University of Alberta Library
0 1620 2533573 6

Ex libris
universitatis
albertaensis

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Laura Clark

TITLE OF THESIS: The Stanford-Binet Intelligence Scale: Fourth Edition: A Validity Study.

DEGREE: Master of Education

YEAR THIS DEGREE GRANTED: Spring 1988

Permission is hereby granted to **THE UNIVERSITY OF ALBERTA LIBRARY** to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the

Date:

88-34

THE UNIVERSITY OF ALBERTA
FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled THE STANFORD-BINET INTELLIGENCE SCALE: FOURTH EDITION: A VALIDITY STUDY, Submitted by LAURA CLARK in partial fulfillment of the requirements for the degree of Master of Education.

Date: March 8, 1988

Acknowledgements

I would like to express my sincere appreciation to the many individuals who contributed to the completion of this thesis.

To Dr. Henry Janzen, I extend special thanks for his consistent willingness to answer questions, provide clarity and direction and give time and energy to the completion on this project.

Thanks to Dr. John Paterson for his encouragement and optimism throughout and for his helpful feedback. Special thanks for his willingness to serve on my committee.

I would also like to thank Dr. Murray Smith for his willingness to serve on my committee and for his encouraging comments.

Thanks to Hau Chau for her invaluable assistance in the data analysis, to Dan Westfall for his interpretation of the behaviors of the computer, to Bruce Fritzke and Rivvy Meloff for being willing to give assistance and extra resources during this project. I wish to acknowledge the 524 students who collected the data for this project.

Thanks to my family and friends for their support and willingness to listen.

Thanks to Randy Clark for his support and endurance during the long process.

TABLE OF CONTENTS

CHAPTER		PAGE
I	INTRODUCTION	1
	Background to the Study	1
	Purpose of the Study	2
	Sample	3
	Research Questions	3
	Delimitations	4
	Limitations	4
	Overview of the Study	4
II	REVIEW OF THE LITERATURE	5
	History of the Binet and Wechsler Intelligence Tests	5
	Current Literature	8
	Normally Achieving Populations	16
	Current SBIV Studies	16
	WISC-R Pattern Analysis	18
	Learning Disabled Populations	19
	Current SBIV Studies	19
	WISC-R Pattern Analysis	20
	Discrimination of Learning Disabled Samples with the WRAT-R, VMI and PASS	23
	Mentally Retarded Populations	25
	Current SBIV Studies	25
	WISC-R Pattern Analysis	26
	Discrimination of Mentally Retarded Samples with the WRAT-R, VMI and PASS	27
	Gifted Populations	28
	Current SBIV Studies	28
	WISC-R Pattern Analysis	29
	Discrimination of Gifted Samples with the WRAT-R, VMI and PASS	30
	Current SBIV Studies with Preschool Children	30

CHAPTER ..... PAGE
Convergent Validity of the SBIV with the WRAT-R, ...
VMI, and PASS 32
WRAT-R 32
VMI 34
PASS 34
Summary 35
III RESEARCH DESIGN AND PROCEDURES 38
Procedure 38
Data Collection 38
Education Clinic 39
Sample 39
Tests Administered 44
Stanford-Binet Intelligence Scale: Fourth Edition 44
Theoretical Model 44
Subtest Descriptions 45
Reliability 47
Validity 48
Wechsler Intelligence Scale for Children-Revised 49
Description 49
Reliability 50
Validity 50
Wide Range Achievement Test-Revised 51
Description 51
Reliabillity and Validity 51
The Developmental Test of Visual-Motor Integration 52
Description 52
Reliability 52
Validity 52
Perception of Ability Scale for Students 53
Description 53
Reliability 53
Validity 53
Analysis of the Data 54

Missing Data 54
Research Questions and Hypotheses 56
CHAPTER ..... PAGE
IV RESULTS 59
Descriptive Sample Statistics 59
Discriminative Ability of the SBIV, WRAT-R, VMI and ....
PASS for Special Education Groups 62
Research Question 1 62
Hypothesis 1 64
Hypothesis 2 67
Convergent Validity of the SBIV 73
Hypothesis 3 76
Hypothesis 4 76
Hypothesis 5 76
Hypothesis 6 76
Hypothesis 7 77
Hypothesis 8 77
Hypothesis 9 77
Hypothesis 10 77
Hypothesis 11 78
Hypothesis 12 78
Hypothesis 13 78
Hypothesis 14 79
Preschool Profile Analysis 79
Research Question 2 79
Summary 81
V DISCUSSION AND CONCLUSIONS 83
Introduction 83
Profile Analysis of Special Populations 83
Learning Disabled 83
Mentally Handicapped 87
Gifted 87
Normally Achieving 88
Analysis of Variance Findings 90

CHAPTER ..... PAGE
SBIV 90
WRAT-R 94
PASS 94
VMI 95
Group Classification by Linear Regression and .....
Discriminant Function Procedure 95
Convergent Validity 98
Stanford-Binet vs. WISC-R 98
WRAT-R 98
PASS 99
VMI 99
Preschool Group 99
Future Research 100
Summary 101
REFERENCES 104
APPENDIX A: Distribution of the Sample by Demographic.....
Characteristics 113
APPENDIX B: Psychological Testing Center, Faculty of .....
Education Clinical Services: Report Forwarding
Form 118
APPENDIX C: Oneway Analysis of Variance: SBIV, WRAT-R,
PASS and VMI Scores 120

LIST OF TABLES

TABLE		PAGE
3.1	Distribution of the Sample by Tests Given	40
3.2	Reason for Referral by Subgroup	44
3.3	Percentage of Missing Data	56
4.1	Total Sample Frequencies, Means and Standard Deviations for Stanford-Binet Fourth Edition and WISC-R Variables	60
4.2	Total Sample Frequencies, Means and Standard Deviations for WRAT-R, VMI and PASS Variables	61
4.3	Group Means and Standard Deviations for SBIV Subtests	63
4.4	High and Low SBIV Subtest Scores for Learning Disabled, Mentally Handicapped, Normal and Gifted Groups	63
4.5	Scheffe Tests of Comparison for SBIV Composite IQ and Area Scores	65
4.6	Scheffe Tests of Comparison on SBIV Subtest Scores	65
4.7	Scheffe Tests of Comparison on WRAT-R, VMI and PASS Scores	66
4.8	Linear Multiple Stepwise Regression for SBIV Subtests	68
4.9	Linear Discriminant Function of Vocabulary, Number Series, Matrices and Memory for Digits for Prediction of Educational Group Membership	69
4.10	Linear Discriminant Function of SBIV Subtests for Prediction of Educational Group Membership	70
4.11	Linear Discriminant Function of the Stanford-Binet General Abbreviated Battery for Prediction of Group Membership	70
4.12	Linear Discriminant Function of the Stanford-Binet Area Scores for Prediction of Group Membership	71
4.13	Correlations Matrix of SBIV Subtest, Area Scores and Composite IQ with the WISC-R Subtest scores and IQ Scores	74
4.14	Correlation Matrix of SBIV Subtests and WRAT-R Subtests	

PASS Subtests and VMI. 75

TABLE ..... PAGE

I Distribution of the Sample by Age 114
II Distribution of the Sample by Gender 114
III Distribution of the Sample by Ethnicity 115
IV Distribution of the Sample by Locale 115
V Distribution of the Sample by Parental Occupation 116
VI Distribution of the Sample by Parental Education 116
VII Distribution of the Sample by Parental Marital Status 117
VIII Report Forwarding Form 119
IX Oneway Analysis of Variance SBIV Composite IQ 121
X Oneway Analysis of Variance SBIV Verbal Reasoning IQ 121
XI Oneway Analysis of Variance SBIV Abstract/Visual Reasoning IQ 121
XII Oneway Analysis of Variance SBIV Quantitative Reasoning IQ 122
XIII Oneway Analysis of Variance SBIV Short-Term Memory IQ 122
XIV Oneway Analysis of Variance SBIV Vocabulary 122
XV Oneway Analysis of Variance SBIV Comprehension 123
XVI Oneway Analysis of Variance SBIV Absurdities 123
XVII Oneway Analysis of Variance SBIV Pattern Analysis 123
XVIII Oneway Analysis of Variance SBIV Copying 124
XIX Oneway Analysis of Variance SBIV Matrices 124
XX Oneway Analysis of Variance SBIV Number Series 124
XXI Oneway Analysis of Variance SBIV Bead Memory 125
XXII Oneway Analysis of Variance SBIV Memory for Sentences 125
XXIII Oneway Analysis of Variance SBIV Memory for Digits 125
XXIV Oneway Analysis of Variance SBIV Memory for Objects 126
XXV Oneway Analysis of Variance WRAT-R Reading 126
XXVI Oneway Analysis of Variance WRAT-R Spelling 126
XXVII Oneway Analysis of Variance WRAT-R Arithmetic 127
XXVIII Oneway Analysis of Variance PASS Full Scale 127
XXIX Oneway Analysis of Variance PASS General Ability 127

XXX	Oneway Analysis of Variance PASS Arithmetic	128
XXXI	Oneway Analysis of Variance PASS School Satisfaction...	128
XXXII	Oneway Analysis of Variance PASS Reading/Spelling.....	128
XXXIII	Oneway Analysis of Variance PASS Penmanship/Neatness	129
XXXIV	Oneway Analysis of Variance PASS Confidence	129
XXXV	Oneway Analysis of Variance VMI.....	129

Chapter One
Introduction

Background

The Stanford-Binet Intelligence Scale was recently revised (Thorndike, Hagen and Sattler, 1986a). The Stanford-Binet Fourth Edition (SBIV) was divided into four main subject areas which included Verbal Reasoning, Abstract/Visual Reasoning, Quantitative Reasoning and Short-Term Memory. Each of these subject areas included three or four different subtests. The subtest scores within each reasoning area were combined to form a Standard Age Score (SAS). The SAS scores were combined to form a Composite Score.

The design for the standardization of the SBIV was based on five variables: geographic region, community size, ethnic group, age and gender. Data from the 1980 U. S. census were used to allocate sample proportions for these five variables. The standardization sample included 5,013 people. There were 17 age groups with one year divisions from age 2 to 17, and an adult group which included subjects from 18 to 23 years of age. Each age group included 200 to 300 examinees.

The Technical Manual is the primary source of data pertaining to the SBIV (Thorndike, Hagen and Sattler, 1986b). A few empirical studies related to the SBIV have been published (Carvajal and Weyand, 1986; Rothlisberg, 1987). A study by Meloff (1987) provided Canadian norms for the SBIV and suggested that the new Binet was able to differentiate learning disabled from mentally retarded populations and gifted from normal populations. Furthermore, the convergent validity of the SBIV was demonstrated using comparisons with achievement and personality measures. Several commentary articles have also been written (Barnard, 1987; Fagan, 1986; Keith, 1987; Osberg, 1986; Sandovil and Irvin, 1986; Schakel, 1986; Sim, 1987; Slate, 1986; Telzrow, 1987; Vernon, 1987). Further research was needed to substantiate the convergent and diagnostic validity of the SBIV and to address some of the criticisms which have been made in recent articles.

Purpose of the Study

The recent revision of the Stanford-Binet has attracted interest from the educational and psychological communities. The test scores were useful as clinical information, but more research was needed to demonstrate its technical adequacy. In general, the intent of this study was to establish the efficacy of the Stanford-Binet Fourth Edition.

The use of the SBIV is questionable unless it correlates highly with similar intellectual measures, such as the Wechsler Intelligence Scale for Children-Revised (WISC-R). The SBIV has been criticized for being too highly priced in comparison with the WISC-R (Slate, 1986). If the Stanford-Binet is comparable to the WISC-R and offers extra advantages to the tester, extra cost is justifiable. The SBIV is normed for use with a wide age range and may be more cost efficient. There is data in the Technical Manual as to the relationship of the WISC-R and the SBIV, but more research is necessary.

Information on the convergent validity of the SBIV was required to establish adequacy. Results of the study by Meloff (1987) indicated that the SBIV correlated highly with an achievement measure (the Wide Range Achievement Test- Revised) and moderately with visual motor tasks (Developmental Test of Visual Motor Integration) and an academic self concept measure (Perception of Ability Scale for Students). The present study was conducted to substantiate these findings.

The diagnostic utility of the SBIV was an important area to investigate. The authors of the SBIV claimed that this test could discriminate between the three exceptional populations of the learning disabled, mentally retarded and gifted populations. Meloff (1987) has confirmed these claims. However, there is no information on what patterns of scores emerge for special education groups on the SBIV and related psychological and educational measures. In this study, subtest profiles of these groups were analyzed in detail. High and low subtest scores for each group were recorded and compared. Various subtest configurations were analysed to determine which would be the most accurate predictors of membership in a particular educational group.

The authors' of the Fourth Edition state that one of the test's major strengths is that it provides a tool for assessing cognitive development from

preschool years (age two) to adult years (age 23). This study also overviewed the typical pattern of high and low scores for a preschool group.

Sample

The sample for this study was comprised of clinic referrals from the Education Clinic at the University of Alberta. Clinic referrals are received from schools, medical settings, family members or the individuals, themselves. All referrals are voluntary.

Research Questions

The following research questions formed the basis of the study.

1. What subtest patterns of high and low scores are found for the learning disabled, mentally handicapped, normally achieving and gifted groups on the SBIV? Are these subtest patterns similar to those found in the literature for the WISC-R?
2. Are there significant differences on the mean scores of the Binet IV subtests, factor areas and Composite IQ for learning disabled, mentally handicapped, gifted and normal groups?
3. Are the mean scores on the Wide Range Achievement Test - Revised (WRAT-R), Developmental Test of Visual-Motor Integration (VMI) and Perception of Ability Scale for Students (PASS) significantly different for learning disabled, mentally handicapped, gifted and normal groups?
4. Which subtests are the best predictors of membership in the learning disabled, mentally handicapped, gifted or normally achieving groups?
5. What is the relationship between the SBIV and the WISC-R? Are there significant correlations between the Binet Composite and WISC-R Full Scale; the Area scores and the Verbal and Performance scales; and related subtests on both measures?
6. What are the correlations between other psychological and educational measures (WRAT-R, VMI and PASS) and the SBIV?
7. What subtest patterns of high and low scores are found for Preschool children on the SBIV. What are the mean scores on the Composite and factor areas?

Delimitations

This study focused on four main areas.

1. The diagnostic utility of the SBIV and other measures with learning disabled, gifted mentally handicapped and normal populations was investigated.
2. The relationship between the SBIV and a similar intellectual measure, the WISC-R was determined.
3. Evidence for convergent validity of the SBIV was provided. Comparisons with the WRAT-R, VMI and PASS were made.
4. The subtest, area, and composite scores on the SBIV for the preschool group were analysed to find developmental trends.

The study was descriptive in nature. The theoretical basis of the test was not discussed at great length. Analysis focused on mean test scores obtained by specific groups and any significant differences in these scores. An investigation of the underlying structure of the test, using factor analysis, was not attempted in this study.

Limitations of the Study

1. Subjects were clinic referrals and therefore generalizations to the general population should be considered cautiously.
2. The use of more than one experimenter may account for individual differences in the administration and scoring on all tests given.

Overview of the Study

In Chapter I, a description of the new Stanford-Binet Fourth Edition was given and the purpose of this study was presented. In Chapter II a review of the related literature was provided. In chapter III the design of the study was outlined. This description included a description of the sample, description of the instruments used, and an explanation of the data collection and techniques of analysis. Chapter IV contained an analysis of the obtained data. In Chapter V the results were discussed with conclusions and implications given.

Chapter Two
Review of the Literature

History of the Binet and Wechsler Intelligence Tests

The tests which Wechsler and Binet developed continue to be among the 15 tests most frequently used in North America (Lubin, Larsen and Matarazzo, 1984). An understanding of the history of these two scales provides information on the strengths and weaknesses of these respective scales. Each new revision of a test, such as the Binet, attempts to improve the validity and reliability of a test. An outline of the development of the Binet intelligence test serves to underscore the improved features of the Fourth Edition.

The first Wechsler scales was published in 1939. The primary purpose of the Wechsler-Bellevue Intelligence Scale was to provide an intelligence measure for adults. Wechsler deemed the previously developed intelligence scales to be unsuitable for adults. The test items in previous scales had been created for children and therefore were not of interest for adults. There was an overemphasis on speed in previous intelligence scales. Furthermore mental age norms, as included in the Binet tests, were not applicable to adults. As a final criticism, Wechsler noted that few adults were included in the standardization samples of previous tests (Anastasi, 1982). The Wechsler Bellevue was replaced in 1955 by the Wechsler Adult Intelligence Scale (WAIS). The WAIS was very similar in form and content, with some changes in the standardization sample. The reliability of subtests was also increased with the new modification. The latest revision of the WAIS is the Wechsler Adult Intelligence Scale-Revised (WAIS-R) which was published in 1981. The WAIS-R includes Performance and Verbal scales. These two scales consist of five and six subtests, respectively. Administration of the test yields individual subscale scores, a Verbal IQ, a Performance IQ and a Full Scale IQ. The WAIS-R can be used for adults aged 16 years and older.

The primary intent in the publication of the Wechsler-Bellevue was to provide an appropriate intelligence test for adults. In 1949, the potential age range of the Wechsler scales was increased and an instrument for children, the Wechsler Intelligence Scale for Children (WISC), was published. This test was first prepared as a downward extension of the Wechsler-Bellevue. Many

items were taken directly from the adult scale; easier items were added to each subtest. The 1974 revision, the Wechsler Intelligence Scale for Children-Revised (WISC-R) was based on a new standardization sample. Minor changes in item content, scoring and administration were done. The Verbal Scale and the Performance Scale on the WISC-R include six subtests each, two of which are supplementary tests. Ten of the individual subtest scores are used to calculate the Verbal IQ, the Performance IQ and the Full Scale IQ. The test is appropriate for children from six years of age to sixteen years of age. The Wechsler Preschool and Primary Scale of Intelligence (WPPSI) was published in 1967. It can be administered to children from four years to six and a half years.

The development of the Stanford-Binet scale began in the early 1900's. In 1904, the Minister of Public Instruction in Paris formed a committee to find a method to differentiate between mentally deficient and normal children. The purpose of the 1905 scale, which Alfred Binet developed in collaboration with Theodore Simon, was to assess the intellectual ability of school children (Sattler,1982). Thirty tests of mental functions were developed. These tests tapped skills in the areas of perception, auditory and visual memory, word knowledge, social functioning, abstract spatial reasoning and language development (Sattler, 1982).

The 1908 revision of the Binet-Simon scale contained the first age scale. Each item was placed at an age level where 60 to 90 percent of normal age level children passed it. The concept of mental age was introduced into this test. Mental age can be defined as an index of mental performance. It is the "degree of general mental ability possessed by the average child of a chronological age corresponding to that expressed by the MA score" (Sattler, 1982, p. 639). Credit in months was given for each item passed on the Binet. The final score was calculated to provide an overall mental age. One of Binet's assumptions was that a person was normal if performance was equal to other children of their own age; deficient if test performance was equivalent to that of younger children and accelerated if performance level exceeded that of persons of equal age (Terman and Merrill, 1973).

The test was again revised in 1911 at which time some content was changed. The age range covered by this test extended from three years to the adult level. In 1916, the Binet-Simon was revised by Lewis Terman in the

United States. From Binet's original 54 subtests the number of tests was increased to 90. The entire scale was restandardized using an American sample. The concept of IQ was now utilized. Terman based this idea on Stern's mental quotient which is calculated by dividing mental age by chronological age. Terman and his colleagues renamed the ratio, calling it the intelligence quotient. They also added detailed instructions for administering and scoring each test. The resultant new version was called the Stanford Revision of the Binet-Simon Intelligence Scale.

The second revision of The Stanford-Binet was published in 1937 (Terman and Merrill, 1937). This revision consisted of two equivalent forms, L and M. Restandardization took place with a U.S. sample which was more extensive for each age group. However, the sample was proportionately higher than the U.S. population in socioeconomic level and in urban representation. The sample included only American born whites. In this revision, new types of tests were added to the preschool and adult levels; the scale was extended downward to year level II and upward to the Superior Adult III Level. Further statistical analysis was done on all test items to ascertain their validity for inclusion.

A third revision was published in 1960 (Terman & Merrill, 1960). In this version the best items from the two 1937 forms were incorporated into a single form. Restandardization was not completed. Instead, items were selected from forms L and M on the basis of performance of 4,498 examinees who had taken the test between 1950 and 1954. The 1937 scale's IQ tables were replaced by Pinneau's Deviation IQ tables. The Deviation IQ is a normalized standard score with a mean of 100 and a standard deviation of 16.

In 1972, Form L-M was restandardized using a new sample of approximately 2100 cases, with 100 subjects at each Stanford-Binet year level. Subjects for the SBLM were siblings of those chosen for the standardization sample of the Cognitive Abilities Test. This sample was chosen from communities stratified in terms of size, geographical region, and economic status. The sample included black, Mexican-American and Puerto Rican children.

The 1986 revision of the Stanford-Binet Fourth Edition was based on the largest standardization sample of any individualized intelligence test (Osberg, 1986). The authors' described two major strengths of this new test. It has an adaptive-testing format in that items are given based on individual

capabilities. The Fourth Edition also provides a continuous scale for appraising cognitive development from age 2 to adult. The manual also claimed that this new instrument can discriminate between special educational groups. Preliminary information was given in the manual on the relationship of the Binet with other intelligence tests. The purpose of this study was to investigate further the diagnostic utility of the SBIV and its convergent validity.

Current Literature on the Stanford Binet Fourth Edition

The Stanford-Binet Fourth Edition has attracted the interest of psychometricians and psychologists. To date, most published information is in the form of critiques or commentaries. Initially, there was some controversy as to whether or not the new Binet revision should be utilized by psychologists. At a July, 1986 meeting of the Executive Board of the National Association of School Psychologists (NASP) passed a resolution that the Stanford-Binet not be used.

> The National Association of School Psychologists acknowledges concern over issue recently presented to it for the Stanford-Binet IV. The lack of validity data, a technical manual, and the problems of distinguishing the original manual from the "second printing" (which contains the corrected norm tables) are considered to be serious hindrances to the approximate use of that scale. (Fagan, 1986a, p. 1).

The NASP stated that they were not questioning the technical adequacy of the test. In September 1986, the NASP reconsidered and withdrew the July resolution. It was determined that the publisher had taken the required steps to distribute technical manuals, and ensure that corrected manuals were available to purchasers of the test (Fagan, 1986b) .

Further criticism of the SBIV was found in The Alberta School Psychologist. Barnard (1987) stated that the authors of the test have not provided necessary data to verify the technical adequacy of the test. He recommended that the test not be used for decision making until further information was provided. Specifically, he stated that evaluation of the test should examine four critical areas: "measurement validity, examinee

appropriateness (test format), administrative usability, and normed technical excellence (test reliability and range of coverage)." Some of these requirements have been fulfilled through data in the Technical Manual.

The SBIV has been criticized in a recent article by Slate (1986). Slate discussed several perceived weaknesses of the new edition.

1. Slate criticized the publishers for failing to provide sufficient data on the new instrument. For example, the standard error of measurement was not available at the time when Slate published his article. Sandovil and Irvin (1986) have also criticized the SBIV for providing limited information on reliability and validity.

2. The percentage of "g" which the revised Stanford-Binet measures was not explicitly stated. Slate suggested that if "g" accounts for the same amount of variance as has been found on the WISC-R (about 82%, Kaufman, 1975), then the four factor areas of the SBIV would not actually be measured. The four areas of Visual Reasoning, Abstract/Visual Reasoning, Quantitative Reasoning and Short-Term memory may not be appropriate divisions if the variance attributable to "g" is too high.

3. The standardization sample for the SBIV does not reflect national norms in educational and occupational groups. The Stanford-Binet's sample of college graduates was 36.7% while the national population includes only 21.8% at this occupational level. Slate asserted that the weighting procedures which are included in the manual do not alleviate this problem. Sandovil and Irvin (1986) also stated that the weighting procedure is an unsuitable methodology.

4. This edition has retained a standard deviation of 16 when most other intelligence, achievement and psychometric instruments use a standard deviation of 15. Slate suggested that changing the deviation score to 15 would be more beneficial to future users of the test.

5. The cost of the SBIV is over twice that of the Wechsler Adult Intelligence Scale-Revised. Slate recommends the continued use of the Wechsler scale.

6. Different examiners may use different sets of subtests to obtain an IQ score. Slate queried how one can determine the true score if different sets of tests are used to obtain the scores. He cited a hypothetical example in which a school psychologist administers four tests and a private psychologist gives four other tests and two different scores are obtained. Vernon (1987) has also commented on this problem.

7. Factor analytic data has shown that several subtests do not load on their respective factors. For example, Bead Memory has the smallest loading of any of the memory subtests and does not load on the Short-Term Memory factor except for the older age groups.

8. The Stanford-Binet is based on a three-level hierarchical model of intelligence. At the highest level is a general reasoning factor or "g." The second level is comprised of Cattell-Horn's Crystallized abilities, Fluid-Analytic abilities and a Short-Term Memory factor. The third level contains more specific factors of verbal reasoning, quantitative reasoning and abstract/visual reasoning. Slate commented that this theory is only partially supported by factor analysis. Sandovil and Irvin (1986) and Vernon (1987) have also questioned the validity of the hierarchical model.

Sim (1987) and Osberg (1986) did not agree with all of Slate's criticisms and provided counterpoints to his arguments. The first concern, a lack of essential data, was effectively answered by the provision of technical manuals to all purchasers. In particular, information about the standard error of measurement is provided by age for subtest, area and composite scores in both the Guide (pp. 146-50 and 189-91) and the Manual (pp.39-44) (Osberg, 1986; Sim, 1987).

Secondly, the amount of variance which the general factor accounts for is discussed in the manual. Evidence for this is found in the Technical Manual showing that the first common factor extracted accounts for 42 to 50 percent of the total variance. The remaining factors show limited variance, accounting for 2.2 to 9 percent of the variance, depending on which age group is assessed. Specific variance, or information not accounted for in the first four factors obtained is as high as 30 percent, depending on the specific age group which is analyzed (Osberg, 1986 ; Sim, 1987).

In regard to norming procedures, Osberg stated that the proportional weighting procedures were adequate to adjust the sample composition to reflect national proportions of occupational and educational levels. Furthermore, it is common practice to use such weighting procedures and these statistical techniques are fully described in the Technical Manual. Sim was more tentative and suggested that test users should be aware of the demographic characteristics of their own area.

Osberg maintained that the use of a standard deviation of 16, instead of 15, was a minor issue. The Stanford-Binet has always used a deviation of 16

points. He argued that assigning the same means and standard deviations to different tests might encourage the concept that all intelligence tests measure identical construsts. However, for those who wish to use it, a table comparing scores with a mean of 100 and standard deviations of 15 or 16 is provided in the Technical Manual.

The cost of the new Stanford-Binet cannot be directly compared to a similar Wechsler instrument without considering the needs of the psychologist. As both Sim and Osberg noted, the Fourth Edition can be used for clients from the preschool to adult years. The use of Wechsler tests would require at least two different forms and possibly three to cover equivalent age ranges. Each purchaser must consider his own needs. For those who work with a large age range, the Binet may be more cost effective (Vernon, 1987).

The problem of inter-examiner differences using two different sets of subtests was addressed by Osberg. He argued that a "true" score is not possible to achieve since all scores obtained from test batteries are estimates. Osberg did agree that the use of abbreviated batteries would result in differences in individual scores. The full battery should be given to provide the most reliable estimate. Osberg suggested that when the test scores are to be used for educational decisions, all subtests should be given or the relevant organization should agree before hand as to which tests should be a part of a battery.

Slate stated that Bead Memory has the smallest loading of all memory tests on the memory factor. Osberg noted that the other three memory tests - Memory for sentences, Memory for Digits and Memory for Objects, all measure sequential short-term memory. Osberg argued that Bead Memory measures simultaneous memory skills. Based on this apparent successive/simultaneous division, he contended that the respective difference in factor loading is to be expected.

Finally, Osberg considered the hierarchical model of the SBIV. Osberg did not discuss any specific factor analytic data. He simply stated that "the purpose of the model was only to provide for the authors a conceptual framework for test development and construction" (p. 3).

Vernon (1987) has recently reviewed the Stanford-Binet Fourth Edition and has discussed both strengths and weaknesses in regards to administration, scoring and test format. Issues related to technical adequacy and clinical utility were also considered.

A major improvement over the old Binet is the ease of administration for individual items. Most items are self-contained in bound books, with the instructions clearly visible to the examiner as each task is being completed. Furthermore, the scoring is straightforward, although some study and practice is needed to learn procedures. The administration time is relatively short. The manual states that total time for the test is a maximum of 75 minutes. Vernon stated that 90 minutes is a more accurate estimate. Abbreviated batteries can be used which further decreases the time. However, Vernon stated that increased flexibility may result in decreased inter-rater reliability.

One main disadvantage of the new edition is the amount of time required to learn administration and scoring procedures. Furthermore, it requires more time to train students in administration of the test Vernon also described the new test format as lacking the clinical flexibility of the old Binet. Vernon asserted that this test should not be called a Stanford-Binet scale since it is more similar in structure and content to the Wechler scales.

Consideration of the factor analysis of the Fourth Edition caused Vernon to question the validity of the four group factors. Vernon calculated the total group factor content to be 11%. The use of four group factors or areas besides *g* is not justified, in the opinion of Vernon. The inclusion of the Memory area seems especially questionable. The loading of the Memory area on the General Ability factor is only .42. The Bead Memory test has a mean loading on the Memory factor of only .12. The use of the Short-Term Memory Area is unjustifiable because of weak factor analytic support. Vernon also commented that this area may be misinterpreted by teachers and psychologists as representing the memory processes which are used in school learning. Based on these aforementioned data, Vernon concluded that the hierarchical theory of the SBIV is inadequately confirmed.

Vernon discussed several other technical details of the new Binet. These comments are based on a review of data from the manual. The internal consistency of all subtests were high for all age levels. Nearly half of the total reliability coefficients reached .90 and over. Only Memory for Objects was relatively weak. This may be have been due to the small number of items included in this subtest, which totalled only 14. Test-retest reliability coefficients were much lower for individual subtests. The reliability coefficients for the subtests were obtained by analyzing the scores of a small sample of 5 and 8 year olds. For the group of 5 year olds the reliability

coefficients ranged from .56 (Bead Memory) to .78 (Memory for sentences) and were all below .80. In the older group, the reliability coefficients ranged from .28 (Quantitative) to .86 (Comprehension). Most of the subtests had reliability coeffiients below .80. The reliability coefficients for the composite scores were .90 and .91.

Data presented in the manual presented confirmation of the convergent validity of the SBIV. Correlations of area and composite scores were given in the manual with the Stanford-Binet: form L-M (SB-LM), the WPPSI, WISC-R, WAIS-R and the Kaufman Assessment Battery for Children (K-ABC). The mean IQ correlation of all these tests with the SBIV was .85. Vernon observed that this does not mean that any of the tests are interchangeable with the SBIV. Difference in IQs on any two tests might exceed +/- 10 points. Keith (1987) and Sandovil and Irvin (1986) have also noted that the SBIV demonstrates impressive reliability and validity.

The sampling procedures for the standardization of the Fourth Edition are adequate. For most age groups the sample size included 200 to 400 subjects. The sample is reasonably representative of the U.S. population. However, Vernon stated that the corrections for the large proportion of college graduates in the oldest group of 18 years and older were not effective. He did not regard the older group as representative of the U.S. population and suggested that it be discarded. There are some problems in the sampling for blacks and the Asian/Pacific group. There were 14.4% of Blacks in the standardization sample for the SBIV. The percentage of Blacks in the U.S. population is 11.5. The Asian/Pacific group yielded very different scores for different age groups. For example, the mean composite score for the 2 to 6 year group was 88.0 while for the 7 through 22 year group, it was 103.6. These two groups do not appear to be homogeneous.

Vernon also discussed the clinical and diagnostic utility of the SBIV with special education and developmental groups. The composite scores described in the manual for the gifted group are similar to those obtained on the SBLM and the WISC-R. However, Vernon stated that the utility of the composite score for gifted groups is still uncertain. Number series may serve as a useful diagnostic subtest. The Memory subtests are described as valueless for this special education group. Composite IQ means for learning disabled groups are similar to those obtained on the SBLM, WISC-R or K-ABC. The Binet IV appears to have a lower floor for mentally retarded groups.

These results suggest that the Binet is a valid instrument for work with exceptional children.

The SBIV can be used for a wide age range, from age two to adult. The content of the test may vary at different ages. Younger children are given only 8 of the 15 subtests and very few individuals are administered all 15 subtests. This was also a weakness of the SBLM. Vernon argues that alterations in content are required to match developmental changes. For children and adolescents, the SBIV is an adequate tool. However, for use with adults, Vernon noted two problems. First of all, there are no reliable norms for adults older than the age of 23 years. Secondly, administration time for adults may be excessively long.

There is some cultural bias in the new SBIV. Several items are American in content and handicap test takers from other English speaking countries. The Quantitative and Equation Building subtests are particularly guilty of bias. For Canadians, no allowance has been made in the Arithmetic test for conversion to metric units of measurement and item 8 in the vocabulary test requires identification of the American flag.

Vernon also analyzed the content of the SBIV. The four Verbal subtests have a good sampling of skills. The Absurdities test could be improved by adding verbal, as well as pictorial content. The Verbal Relations items are described as being more "artificial" than the Wechsler Similarities subtest items. In comparison with the SB-LM, the new Binet shows less variation in the sampling of verbal abilities. Vernon notes that in the SBIV there are no rhymes, analogies, sentence completions, ideational fluency, interpreting of proverbs, comprehension of paragraphs or verbal reasoning problems. The Abstract/Visual Reasoning area includes subtests which are heterogeneous in nature. These may be superior to the Wechsler Performance subtests, but less extensive than the Kaufman Simultaneous Processing Battery. The Quantitative area is made up of three tests. Vernon states that none of the three appear to involve much mathematical reasoning. The Short-Term Memory area includes four quite varied subtests. These all involve reproduction of nonmeaningful stimuli. Their intercorrelations are somewhat lower than those in the other three areas. This suggests that each subtest shows high specificity.

There are a few subtests which Vernon judged should not be given in the SBIV battery, particularly in a time limited situation. He considered these to be

the weakest subtests. These subtests included Absurdities, Paper Folding and Cutting and Equation Building.

Vernon suggested that more research is needed to validate the use of the SBIV. He recommends that Canadian psychologists give the test to a diverse group of subjects with a wide age range. A pooling of obtained data is the proposed method to provide concensus on use of an abbreviated battery or a modified battery.

The recently published critiques of the SBIV have provided summation of its current strengths and weaknesses. Some empirical research on the SBIV has been completed by Meloff (1987). This study has provided data which upholds the SBIV as a valid and reliable test. In general, Meloff's findings were supportive of the data which are found in the technical manual. A comparison of the American norms from the manual and a Canadian sample showed that these two groups obtained similar mean scores but there were higher standard deviation scores for the Canadian sample. Significant differences between the two populations were reported in eight individual subtests. These subtests included Vocabulary, Comprehension, Absurdities, Pattern Analysis, Matrices, Number Series, Memory for Digits and Memory for Objects. Significant differences were reported for Verbal Reasoning IQ and the Composite IQ between the Canadian and American Samples.

Results of Meloff's study also indicated that the SBIV is able to differentiate learning disabled from mentally retarded populations and gifted from normal populations. Meloff's results substantiate claims made in the manual that the SBIV can discriminate between special educational groups. Meloff's study included a description of which SBIV subtests discriminated among the four educational groups of learning disabled, mentally retarded, gifted and normal children. Verbal Relations, Paper Folding and Cutting and Equation Building were not included in this study since very few examinees were administered these subtests. For the remaining twelve Binet IV subtests only the Vocabulary subtest showed significant differences between all four groups. Comprehension, Quantitative, Bead Memory and Memory for Sentences showed significant differences among all four groups except between learning disabled and normal subjects. Absurdities, Pattern Analysis, Copying, Matrices and Memory for Digits indicated no significant differences between the learning disabled and normal groups nor between the gifted and normal groups. However, these five subtests did differentiate

the mentally retarded from the other three educational groups and discriminated between the learning disabled and gifted populations. Number Series differentiated between the LD and MR groups; the normal and LD groups and the gifted and LD groups but not between the gifted and normal populations. Memory for Objects differentiated the mentally retarded group from the other three populations.

Correlations between the Binet IV and other psychological and educational measures provide initial support for convergent validity of the Binet. Meloff reported findings of significantly high correlations between the Binet IV Composite, Area Scores and individual subtests and the WRAT-R. More moderate correlations were reported between the Binet IV and the VMI. Moderate correlations were also found between the SBIV and the PASS.

Normally Achieving Populations

Current SBIV Studies for Normally Achieving Children

Carvajal and Weyand (1986) administered the General Purpose Abbreviated Battery of the Stanford Binet: Fourth Edition and the Wechsler Intelligence Scale for Children-Revised to 23 children (11 boys, 12 girls) in a third-grade class. The Abbreviated Battery has six subtests including Vocabulary, Bead Memory, Quantitative, Memory for Sentences, Pattern Analysis, and Comprehension. The relationship of the Composite standard age scores of the SBIV and the the Full Scale IQ of the WISC-R was significant (r=.78). The correlations of the two vocabulary subtests and the comprehension tests were both significant (r=.65 and .42 respectively). However, the relationships of other specific subtests on the two instruments which appear to be similar were quite low. In particular, Pattern Analysis and Block Design scores correlated .18 whereas Quantitative and Arithmetic correlated .14.

Rothlisberg (1987) compared the scores on the SBIV and WISC-R for a sample of 32 nonexceptional children in grades one and two. The WISC-R Full Scale IQ (112.5) was significantly higher than the SBIV Composite IQ (105.5). Rothlisberg stated that this difference was due to the students' higher scores on the WISC-R Performance scale. Scores on the Copying subtest were significantly lower than the other A/VR subtests and this may have

contributed to the significant difference between the two tests. The WISC-R Full Scale and Composite correlated .77. The correlations of the Verbal and Performance IQ's were .62 and .51 respectively. Intercorrelations of the Area scores with the FSIQ were all significant and ranged from .39 (Quantitative) to .66 (Abstract/Visual Reasoning, Short-Term Memory). All Area scores except for the Quantitative Reasoning Area correlated significantly with the Verbal IQ. The Verbal Reasoning Area showed the highest correlation with the VIQ (.70), followed by Short-Term Memory(.53) and Abstract/Visual Reasoning (.43). The A/VR area correlated the highest with the Performance IQ (.55). Quantitative Reasoning and Short-Term Memory obtained significant correlations with the PIQ (.40 and .41, respectively), but Verbal Reasoning did not correlate significantly (.07). These results provide some support for the predictions of Thorndike et al. (1986b) that Verbal Reasoning, Short-Term Memory and Quantitative Reasoning would correlate higher with the WISC-R Verbal scale and Abstract/Visual Reasoning would correlate higher with the Quantitative scale. Rothlisberg asserted that the Quantitative Reasoning Area showed low correlations with the WISC-R because only one subtest (Quantitative) is given at younger age levels. It was suggested that the correlation coefficient of the SBIV and other intelligence measures may vary in relation to the age of the sample and which SBIV subtests were administered. Rothlisberg recommended that further study be done as to the effect of subtest configuration on the convergent validity of the SBIV.

In the technical manual, the SBIV and WISC-R were compared using a nonexceptional sample. The mean Composite score on the SBIV was 102.4; the mean Full Scale IQ on the WISC-R was 105.2. The Verbal Reasoning, Quantitative Reasoning and Short-Term memory scores were combined for a comparison with the mean Verbal IQ. The two resultant IQs were 103.0 and 103.9, respectively. A comparison of the Abstract/Visual Reasoning score and the Performance mean scores resulted in a difference of 6.4 points. The Abstract/Visual Reasoning (AVR) mean score was 98.9; the WISC-R Performance mean was 105.3. Thorndike and his associates accounted for this difference in that the AVR area on the SBIV involves more abstract inductive reasoning than the WISC-R Performance tasks.

The relationship between the WISC-R and the SBIV was examined in the manual using a correlational analysis. The correlation coefficient for the composite score (SBIV) and WISC-R Verbal IQ was .78; with WISC-R

Performance IQ it was .73; and with the WISC-R Full Scale IQ it was .83. The scores on the Verbal Reasoning Area of the SBIV correlated .72 with WISC-R Verbal scores, .60 with WISC-R Performance scores, and .73 with WISC-R Full Scale scores. This was as expected, since Thorndike and his associates predicted that Verbal Reasoning would correlate more highly with scores on the Verbal Scale and Full Scale of the WISC-R than with the scores on the Performance Scale. They also predicted that Quantitative Reasoning and Short-Term Memory areas would show a similar pattern. Furthermore, they predicted that the Abstract/Visual Reasoning Area scores would correlate more highly with scores on the Performance Scale than with scores on the Verbal scale. These latter expected patterns were not produced. Scores on the Abstract/Visual Reasoning, Quantitative Reasoning and Short-Term Memory Areas correlated about equally as high with the WISC-R Verbal and Performance scores. For the WISC-R Verbal scores, the correlations ranged from .64 to .68, and for the WISC-R Performance scores, they ranged from .63 to .67. The correlations in these three Areas ranged from .69 to .73 with the WISC-R Full Scale IQs.

Meloff (1987) compared Canadian and American samples of normally achieving children. Significant differences were reported for Verbal Reasoning IQ and the Composite IQ between the two samples. The mean Verbal Reasoning IQ was 104.3 for the Canadian sample and 100.1 for the American sample. The Canadian and American Composite IQs were 103.4 and 99.7, respectively. These findings are consistent with the predictions of Vernon. He suggested that since one quarter of the American standardization sample was nonwhite, mean IQs of white, English speaking Canadian children would average considerably over 100, probably 104. In Meloff's study, 89% of the sample were caucasian subjects.

WISC-R Pattern Analysis for the Normally Achieving Population

Mueller, Dash, Matheson and Short (1984) used meta-analytical procedures to study the typical scoring pattern of average children on the WISC-R. They found that normally achieving children demonstrated strengths on perceptual or visual tasks. Average children usually achieved their highest scores on Object Assembly and Picture Arrangement. This relative strength in perceptual tasks was described by Mueller et al. as comparable to the

profile which mentally handicapped subjects usually obtain. The average children in this study achieved weak scores in the Arithmetic, Information and Coding subtests.

The subtest scatter of normal children has been investigated by Kaufman (1976). It has been shown that normal children do not demonstrate a "flat" profile where all scores deviate only slightly from or equal the mean subtest score. Frequently, as many as two out of 12 WISC-R scores may vary by three of more points from the mean subtest score for children in the non exceptional population.

Learning Disabled Populations

Current SBIV Studies with Learning Disabled Samples

The technical manual includes a study which compared the SBIV and the WISC-R scores of learning disabled (LD) children. The mean composite score on the Fourth Edition was 84.8 and the mean of the WISC-R Full Scale IQ was 87.8, a difference of 3 points. The combined mean score of Verbal, Quantitative Reasoning and Short-Term Memory areas when compared to the mean of the WISC-R Verbal IQ (VIQ) were quite similar. The largest difference in mean scores on the two tests was between the Abstract/ Visual Reasoning Area and the Performance Scale score, a difference of 4.6 points. Learning disabled children performed better on the WISC-R Performance Scale than on the WISC-R Verbal Scale. This last finding is consistent with research which suggests that learning disabled children perform relatively well on the visual motor tasks on the WISC-R (Kaufman, 1979; Naglieri, 1980; Naglieri, 1985; Schiff, Kaufman and Kaufman, 1981; Smith, 1978; Yewchuk, 1986).

The Composite Score on the Binet IV correlated .84 with the WISC-R Verbal Scale score, .70 with the Performance Scale score and .87 with the Full Scale IQ. The authors of the manual expected Verbal Reasoning, Quantitative Reasoning and Short-Term Memory scores to have higher correlations with the WISC-R Verbal Scale IQs than the scores on the Abstract/Visual Reasoning Area. The Verbal Reasoning score correlated .83 with the Verbal Scale IQs. However, correlations for the other three areas were fairly equal. The Verbal IQ correlated .69 for Short-Term Memory and Quantitative Reasoning scores and .68 with the A/VR score. As expected, the

scores on Abstract/Visual Reasoning correlated more highly with the WISC-R Performance Scale (.72) than with the WISC-R Verbal Scale and had a higher correlation with the WISC-R Performance Scale than the other Area scores.

In Meloff's (1987) study with a Canadian sample of learning disabled children, the sample means for all area scores and the composite were higher than those found for the American sample. In Meloff's (1987) study learning disabled mean scores were approximately 15 points above those cited in the manual. This corroborates the predictions of Vernon (1987).

WISC-R Pattern Analysis of the Learning Disabled Population

The search for a typical pattern of scores on the WISC-R has been ongoing for those involved in learning disabilities research. The difference between the Verbal IQ (VIQ) and Performance IQ (PIQ) scores has been considered as a possible indice. The supposition in using this discrepancy score is that learning disabled populations show more variability in scores than the normal population. The "scatter" found within a profile has also been measured and related to the presence of learning disabilities. The amount of scatter within a profile is measured by subtracting the lowest subtest scale score from the highest subtest scale score (Kaufman, 1976). The results of studies which have used these two measures of variability are inconclusive. In several recent studies, scaled score ranges and VIQ/PIQ discrepancies obtained in LD samples were not shown to be substantially different from the average group or the standardization sample of the the WISC-R . The use of these types of scores in scatter analysis may have low value as a diagnostic tool, although they may be helpful in individual cases (Dudley-Marling, Kaufman and Tarver, 1981; Galvin, 1981; Kaufman, 1979; Kaufman, 1981; Kavale and Forness, 1984; Leton, Myanoto and Ryckman, 1987; Moore and Wielan, 1981; Naglieri, 1982; Ryckman, 1981; Sattler, 1982; Schiff, Kaufman and Kaufman, 1981; Ullman , Kramer and Schellenberg, 1985;Yewchuk, 1986).

Researchers have investigated the typical high and low scores of LD groups on the subtests of the WISC-R. This type of analysis has shown significant trends. Sattler (1982) reviewed thirty studies of reading disabled children. His results indicated that the the average subtest ordinal ranking from highest to lowest, was the following: Picture Completion, Picture

Arrangement, Block Design, Object Assembly, Similarities, Comprehension, Vocabulary, Coding, Digit Span, Arithmetic and Information. The four most difficult subtests form the acronym ACID (Arithmetic, Coding, Information, and Digit Span). In most of the studies reviewed, the Verbal Scale IQs were lower than Performance Scale IQs. A review by Kaufman (1979) showed a similar rank ordering for learning disabled samples, with a few minor changes. Object Assembly was in the number one position, while Picture Completion, Picture Arrangement and Block Design moved down one step to positions 2, 3 and 4, respectively. Comprehension and Similarities switched positions to become number 5 and 6, respectively. The review by Kaufman did not include the Digit Span subtest.

A recent meta-analysis of 94 studies confirmed that learning disabled groups tend to show a weakness in the ACID tests. Furthermore, relative strengths involve subtests included in the Perceptual-Organization Factor such as Object Assembly, Block Design and Picture Completion (Kavale and Forness, 1984). In summary, these reviews and other current studies indicate that a typical profile for learning disabled samples consists of strengths in perceptual and visual motor areas, a concomitant increased Performance IQ over the Verbal IQ and weaker scores in the ACID subtests (Fowler, 1986; Galvin, 1981; Kaufman and McLean, 1986; McManis, Gigley, Richert, and Fabre, 1978; Moore and Wielan, 1981; Mueller, Dennis and Short, 1986; Naglieri, 1980; Naglieri, 1985; Schiff, Kaufman and Kaufman, 1981; Sexton and Street, 1985; Smith, 1978; Stevenson, 1980; Strichert and Love, 1979; Ysseldyke, Algozzine, Shinn and McGue, 1982). Sattler (1982) cautioned that there is no one particular WISC-R pattern that is reliably diagnostic of learning disability. There are trends which may provide helpful information. Each case must be considered individually.

Factor analytic studies have compared the global scoring patterns of learning disabled children to those of normal children. A summary of research results is given by Sattler (1982). A factor analysis of the WISC-R standardization group revealed that three factors - Verbal Comprehension (VC), Perceptual Organization (PO) and Freedom from Distractibility (FD) were obtained (Kaufman, 1975). The verbal factor included significant loadings by four of the regularly given verbal subtests. Arithmetic has a more moderate loading. Picture Completion and Picture Arrangement also had moderate loadings on the Verbal Comprehension factor, which may indicate that these

two subtests require more verbal mediation than the other Performance subtests. The Perceptual Organization factor seemed to measure a variable common to the Performance scale subtests. Block Design, Object Assembly, and Picture Completion had high loadings on the Perceptual Organization factor, followed by Picture Arrangement and Mazes, with more moderate loadings. These two obtained factors, VC and PO, followed the divisions used by Wechsler and have been equated with the Verbal and Performance areas (Kaufman, 1975; Sattler, 1982). The third obtained factor is generally labelled as the 'Freedom from Distractibility.' factor. In Kaufman's study, the Arithmetic and Digit Span subtests had high loadings on the third factor, followed by Information and Coding B, which had moderate loadings. This third factor can be interpreted as reflecting attentional or concentration skills, or the presence of anxiety. These three factors have been consistently found in studies relating to learning disabilities (Galvin, 1981; Kaufman, 1979; Kaufman, 1981; Kaufman and McLean, 1986; Kavale and Forness, 1984). The Verbal Comprehension and Perceptual Organization factors are the most robust in LD studies. The Freedom from Distractibility Factor is less strong and is considered by some researchers to be negligible (Brock, 1982; Fowler, 1986). Most of the studies on learning disabilities show consistent trends as to the rank ordering of the factors. Perceptual Organization tasks usually represent relative strengths and are the easiest, Verbal Comprehension tasks are less easy and tasks which load on the Third Factor are the most difficult. This can be summarized in the equation FD < VC < P C (Galvin, 1981; Kaufman, 1981; Kaufman and McLean, 1986; Kavale and Forness, 1984; Mueller, Dennis and Short, 1986).

Bannatyne (1974) proposed a recategorization of WISC-R scores. The model was not based on factor analysis, but was based on Bannatyne's clinical judgment as to which subtests measured similar tasks. Subtests were recombined into Spatial, Conceptual, Sequential and Acquired Knowledge categories. The Spatial category included the Picture Completion, Block Design and Object Assembly subtests. The Comprehension, Similarities, and Vocabulary subtests made up the Conceptual category. The Sequential category included a combination of the Arithmetic, Digit Span and Coding subtests. Acquired Knowledge was comprised of Information, Arithmetic and Vocabulary.

Kaufman (1981) compared the first three categories to his factor analytic model and claimed that the labels could easily be interchangeable, that is, the Spatial category can be compared to the Perceptual Organization factor; Conceptual to Verbal Comprehension and Sequencing to Freedom form Distractibility. Acquired Knowledge represented abilities that are usually learned in the home or at school (Sattler, 1982).

Children with learning disabilities frequently have a WISC-R pattern of Spatial > Conceptual > Sequential (Clariozo and Bernard, 1981; Dudley-Marling et. al, 1981; Kaufman, 1979; Kaufman, 1981; Sattler, 1982; Smith, Coleman, Dolecki, and Davis, 1977; Vance and Singer,1979). Acquired Knowledge is generally found to be lower than Spatial and Conceptual categories when it is included (Dudley-Marling et al, 1981; Fisher, Wenck, Schurr and Ellen, 1985; Smith et al, 1977; Vance and Singer, 1979). The Acquired Knowledge category has a high correlation with the Conceptual category and therefore its usefulness as a discriminator has been questioned (Mueller, Dennis and Short, 1986).

The problems with using Bannatyne profiles as a diagnostic scheme have been summarized by Kaufman (1981). Some studies of learning disabled samples have not produced the expected relationships. Furthermore, the proportions of individuals in the group displaying the characteristic pattern generally has been quite small. Kaufman has noted that the Spatial > Conceptual > Sequential pattern has been produced in other groups such as the emotionally impaired, juvenile delinquents and even non-exceptional samples.

In summary, the recent work on profile analysis of LD groups indicated that they show strength in perceptual and visual-spatial tasks, are weaker in verbal skills and are relatively weak in short-term memory or attentional tasks.

Discrimination of Learning Disabled Samples with the WRAT-R, VMI and PASS

In a recent survey of learning disabilities research, 54% of those studies included a standardized achievement measure for identification criteria and the Wide Range Achievement Test (WRAT) was the most commonly used achievement measure (Kavale and Nye, 1981). The popularity of the WRAT instrument is well founded since it has been shown to be highly discriminative

between learning disabilities and normal achievement levels (Janzen, Boersma, Fisk and Chapman, 1983). Koppitz (1981) found achievement scores of an LD group to be about three-and-one-half years below grade level. Prasse, Siewert and Breen (1983) compared the 1978 WRAT scores of LD and normally achieving students and found the WRAT scores of the LD group to be significantly lower. Scull and Brand (1980) found WRAT reading scores to be significantly higher than spelling or mathematics scores. Meloff (1987) reported that the WRAT-R Arithmetic and Reading subtests differentiated between LD, mentally retarded, gifted and normal groups. WRAT-R Spelling scores were significantly different for all groups except the learning disabled and normal subjects.

The Beery Developmental Test of Visual-Motor Integration (VMI) is frequently used to assess visual motor skills. A delay in this area is sometimes present for LD groups. Janzen et al. (1983) used visual-perceptual tests such as the Beery or Bender to discriminate between LD's and normal achievers. Koppitz (1981) found that a group of LD students showed more significant delays in visual-motor perception and were less well organized in written work than a students from a resource room group. However, there were some contradictory findings. Ysseldyke, Algozzine, Shinn and McGue (1982) compared an LD sample to a low achievement sample (who had scored at or below the 25th percentile on an intellectual measure). They found that the VMI was unable to discriminate between the two samples. McManis et al. (1978) also found no significant difference in the scores on visual motor tasks when comparing reading disabled and nondisabled boys. Leton, Miyamoto and Ryckman (1987) performed a cluster analysis on 100 LD boys and found three distinctive subgroups. One of these was characterized by low scores on visual-spatial and motoric subtests. Meloff (1987) investigated the discriminative validity of the VMI and found no significant differences in scores between learning disabled and mentally retarded populations. There were, however, differences between the learning disabled and normal populations on this measure.

Poor achievement in academic subjects due to a learning disability may result in emotional difficulties for students. This assumption has been verified by researchers. Kolligian and Sternberg (1987) summarized the findings:

> Repeated failure experiences lead learning disabled children to believe that they do not have the ability to succeed and that their efforts are therefore futile....Consequently, the learning disabled individual's perception of his or her general competence is, on the average, significantly lower than that of nondisabled children. (Kolligian and Sternberg, 1987, p. 13).

This decreased sense of potency may result in increased emotional problems and/or behavior problems (Schiff et al, 1981; Ysseldyke et al, 1982). A study of academic self concept as measured by the Perception of Ability Scale for Students (PASS) showed that LD children have significantly more negative self perceptions of ability in arithmetic, reading and spelling than do normal control children. The general attitude towards school also appeared to be more negative for LD children (Chapman and Boersma, 1979). However, recent work by Meloff (1987) suggested that the PASS is not able to discriminate between LD children and other special educational groups or normal subjects.

Mentally Handicapped Populations

Current SBIV studies with Mentally Handicapped Groups

In the technical manual is a comparative study of SBIV and WISC-R scores for a sample of Mentally Retarded (MR) students. The highest scores for the MR students were obtained on the Abstract/Visual and Quantitative Reasoning Areas of the SBIV, and on the WISC-R Performance Scale. The lowest score was in the Short-Term Memory Area. The pattern of correlations between the SBIV and the WISC-R was similar to that found for the learning disabled and gifted students. Verbal Reasoning and Quantitative Reasoning area scores showed the highest correlations with WISC-R Verbal IQs. As expected, scores on the Abstract/Visual Reasoning area correlated more highly with the Performance IQ than the Verbal IQ on the WISC-R. The SBIV Composite score correlated .68 with the WISC-R Verbal IQ, .45 with the Performance IQ and .66 with the WISC-R Full Scale score.

Meloff (1987) tested a Canadian sample of mentally retarded students. In this study, all mean scores were higher than those given in the manual with the exception of the Quantitative Reasoning IQ mean. The highest mean

score for the Canadian sample was the AVR area and the lowest score was Short-Term Memory. Meloff's study also showed that the Binet can effectively discriminate between learning disabled and mentally retarded subjects. Significant differences were found between the mentally retarded and learning disabled subjects on all SBIV subtests, Area scores and the Composite IQ scores.

WISC-R Pattern Analysis of the Mentally Retarded Population

The results of profile analysis studies with mentally retarded subjects have suggested that there is relative consistency in the pattern of subtest high and lows scores. In an overview of the literature, Sattler (1982) listed the rank order of subtests from easiest to hardest as follows: Object Assembly, Picture Completion, Block Design, Coding, Similarities, Comprehension, Picture Arrangement, Information , Arithmetic and Vocabulary. There were some changes in the rank order in other studies (Kaufman and Van Hagan, 1977; Naglieri, 1980). The general concensus in the research is that the mentally retarded, as a group, perform best on tasks which require spatial ability and the manipulation of objects and are relatively weak on tasks which tap verbal comprehension and expression (Kaufman,1979; Mueller, Dash, Matheson and Short, 1984; Mueller, Dennis and Short, 1986; Naglieri, 1985). The profile of scores for mentally retarded samples is quite similar to that demonstrated for learning disabled groups. It has been suggested that the WISC-R subtest profile may not distinguish between LD and borderline mentally retarded groups (Naglieri, 1985). It may be that educational groups should be divided on the basis of Full Scale IQ level rather than using a diagnostic cluster (Clariozo and Bernard, 1981; Mueller et al., 1986).

Factor analytic studies have substantiated Kaufman's three factor division for the mentally retarded samples (Cummins and Das, 1980; Kaufman, 1981; Sattler, 1982). It appears that the Freedom from Distractibility factor is less stable for retarded children than for non-exceptional children (Sattler, 1982). Current research suggests that children with low IQs do the best on Perceptual Organization tasks, do less well on Freedom from Distractibility Tasks and show relative weakness in tasks included in the Verbal Comprehension area (Mueller et al., 1984; Mueller et al.,1986). This

can be summarized in the equation Perceptual Organization > Freedom from Distractibility > Verbal Comprehension.

Use of Bannatyne's recategorization system for the mentally retarded population does not show as distinctively ordered patterns as are demonstrated in learning disabled samples. The mentally retarded group tends to score the highest on the Spatial Cluster and lowest on the Acquired Knowledge cluster. There has been less consistency in the obtained rank order of the Conceptual and Sequential Categories (Clariozo and Bernard, 1981; Mueller, Dash, Matheson, Short, 1984; Smith et al, 1977).

These studies have all shown that MR children demonstrate relative strength in visual motor tasks which involve spatial manipulation and visual discrimination. The understanding of verbal relationships and ability to express verbal knowledge is usually a weakness. School related tasks such as arithmetic, also seem to be an area of weakness.

Discrimination of Mentally Retarded Samples with the WRAT-R, VMI and PASS

The Wide Range Achievement Test-Revised (WRAT-R) is used to assess reading, spelling and arithmetic skills . Cummins and Das (1980) investigated the intellectual abilities which might be related to WRAT achievement scores for mentally retarded children. The Freedom from Distractibility Factor score showed the strongest relationship with the WRAT arithmetic subtest score, with a correlation coefficient of .43. The Perceptual Organization Factor score showed a moderate relationship with the Arithmetic subtest score (.22). The Verbal Comprehension factor was unrelated to any of the WRAT subtests. The authors of this study suggest that mentally retarded children may not be applying their verbal abilities to academic tasks. However, it should be noted that mentally retarded children usually obtain the lowest scores on those subtests which are included in the Verbal Comprehension Factor. Therefore, they may rely more on their relatively stronger skills which are tapped by the other two factors.

Meloff (1987) investigated the ability of the WRAT-R, VMI and PASS to discriminate between the mentally retarded population and other special educational groups. Melloff found that the WRAT-R subtests were able to differentiate learning disabled and mentally retarded groups. On the VMI, no

significant differences were noted between the LD and MR populations. On the PASS, the only significant difference was between MR and normal groups on General Ability and Reading/Spelling Scales.

Gifted Populations

Current SBIV studies with Gifted Children

Information is provided in the Technical Manual on comparative studies of the Binet IV and the WISC-R for gifted samples. The mean Composite Score for the SBIV and the mean WISC-R Full Scale IQ were very similar (116.3 and 117.7, respectively). Verbal Reasoning mean scores and Quantitative Reasoning mean scores were averaged and compared to the mean Verbal IQ for the WISC-R. The difference between the obtained scores was minimal (2.3 points) with 115.3 as the Binet mean and 117.6 as the Wechsler score. The largest difference in scores (4.9 points) was between the A/VR area and the Performance area on the WISC-R. Scores on the WISC-R Performance Scale were lower than the Verbal Scale scores. This sample also did less well on the A/VR area than on other areas of the Binet IV.

The manual also provided information on the pattern of correlations found between the two instruments for the gifted sample. The Verbal Reasoning and Quantitative Reasoning Area scores correlated more highly with WISC-R Verbal Scale IQs than did the A/VR and Short-Term Memory scores. Quantitative Reasoning scores and Verbal Reasoning scores both correlated more highly with the WISC-R Verbal Scale IQs than with the Performance Scale IQs. The scores on the A/VR area correlated more highly with the Performance Scale IQs than with the Verbal IQs. The Composite score correlated .62 with the WISC-R Verbal Scale, .62 with the Performance Scale and .69 with the WISC-R Full Scale. The authors stated that the WISC-R scores in this sample are quite restricted in range and therefore they considered these correlations to be satisfactory.

Meloff (1987) compared a Canadian sample of gifted children to the American Sample given in the manual. All Canadian area scores except for Short-Term Memory were approximately one standard deviation higher than the manual scores. As found in the manual, the Canadian gifted sample scored lowest in the A/VR area.

WISC-R Profile Analysis of the Gifted Population

The exact rank ordering of the subtest scores for the gifted varies from study to study. A Meta-analysis of 29 samples of exceptional children by Mueller et al. (1984) indicated that above average children obtained the highest scores on Comprehension and Similarities subtests, middle scores on Object Assembly and Picture Arrangement and the lowest scores on Arithmetic and Coding.

Factor Analytic studies of gifted children have confirmed the factors suggested by Kaufman. The Verbal Comprehension factor and the Perceptual Organization factor are similar to other subgroups. However the third factor may be measuring different skills for this group (Karnes and Brown, 1980; Kaufman, 1981; Sapp, Chissom and Graham, 1985). Karnes and Brown (1980) found that Arithmetic and Picture Completion loaded substantially on the third factor but Coding and Picture Arrangement did not load on this factor at all. Sapp, Chissom and Graham (1985) found information and Vocabulary to load the highest on the third factor . Again, Coding did not load on this factor. The third factor may be different for the gifted sample and must be interpreted with caution.

A sequential ranking of Kaufman's factor scores reveals some information about the relative strengths and weaknesses of above average children. The mean score profile for bright children shows a facility in verbal skills, with the highest mean scores on the Verbal Comprehension factor. Lowest scores are obtained on tasks which require rapid learning, short-term memory, sequential processes and attention or concentration skills. The lowest scores are on the Freedom from Distractibility Factor. Generally, the Perceptual Organization Factor has been ranked second, although it has also been found as the strongest skill (Mueller, Dash, Matheson and Short, 1984; Mueller, Dennis and Short, 1986). In summary, the rank order of factor skills for the gifted from easiest to hardest is best described in the equation: Verbal Comprehension > Perceptual Organization > Freedom from Distractibility.

Information on Bannatyne's recategorization system for the gifted is sparse. However, Mueller et al (1984) suggested that this group obtains the highest scores on the Conceptual cluster and the lowest scores on Sequential tasks.

Discrimination of the Gifted Samples with the WRAT-R, VMI and PASS

As was mentioned in the section of LD subjects, Meloff found that the WRAT-R was able to discriminate between all groups. The Arithmetic, Reading, and Spelling subtests differentiated between the gifted sample and all other groups. The VMI showed no significant differences between the gifted and normal populations in Meloff's study. There were no significant differences for the PASS between the gifted and any other group on the Full Scale score or any individual scales in Meloff's study.

Current SBIV Studies with Preschool Groups

Telzrow (1987) has commented on the utility of the Stanford-Binet for work with preschoolers. The large standardization sample and the special section on preschool assessment in the guide are listed by Telzrow as strengths. The large colorful pictures included in the picture vocabulary section are beneficial for work with young children. However, the lack of practical play material and complexity of instructions have been noted as possible weaknesses of this test (Telzrow, 1987; Vernon, 1987). Telzrow noted that the major disadvantage of the test was the organization of the items by subtest rather than by age levels. It is often difficult to obtain a basal score, especially for younger children at the low end of the IQ distribution. This may result in several estimate scores and cause difficulty in getting a valid composite score. The authors of the guide discussed the problem as follows:

> At age 2, zero scores occur frequently, particularly on tests such as Bead Memory and Quantitative that present novel tasks to the examinee...the frequency of zero scores places a limit on the the lowest composite score that can be obtained on the scale for 2-year-olds. As a result, the Fourth Edition does not discriminate adequately among the lowest 10 to 15% of the 2-year-old group.
>
> At age 3, zero scores occur much less frequently than in the 2-year-old group; however, they do occur. For the 3-year-olds, the Fourth Edition adequately discriminates

among all except the lowest 2% of the age group (Telzrow, 1987, p. 7-8).

Telzrow stated that the Binet IV is limited in its ability to ascertain specific strengths or weaknesses for a preschool child. A maximum of eight subtests may be administered to this age group. The Quantitative area is represented by only one subtest, and the Abstract/Visual Reasoning and Short-Term Memory areas by only two subtests each. Factor Analytic studies showed that no quantitative nor memory factors are present for preschoolers. Furthermore, the area divisions are not confirmed by factor analysis of specific subtests. For example, the Bead Memory subtest loaded highest of all on the Abstract/Visual Reasoning factor for this age group. A valid estimate of "g", verbal factors and visuo-spatial factors appear to be supported for this age group but more specific analysis is not possible.

Schakel (1986) discussed the use of the SBIV on a preschool population. The addition of many nonverbal items and the focus on different types of abilities are cited as advantageous for work with this age group. In particular the separation of a "scholastic" ability factor (crystallized abilities) from other factors is beneficial for work with preschoolers who have no school experience. The adaptive testing format which is described in the manual allows for tailoring of the battery to fit individual needs. The use of the Vocabulary subtest to establish an entry level was also described as an asset of the test, since it minimizes the amount of tests given which are too easy or too hard. In addition, preschoolers show great variation in ability level and using an entry level alleviates some guesswork as to which tasks are appropriate for a specific child. Schakel does add a cautionary note. Research suggests that the predictive validity of traditional IQ tests is greater for children above age 2. Further research is suggested to support the use of the SBIV with preschool children.

The manual included a factor analytical study based on ages 2 to 6. The analysis showed that a general ability factor ("g") was present for this age group. A verbal factor appeared on four tests. A further factor, which was primarily Abstract/Visual in nature, appeared on the other four tests. Only one quantitative test was given and therefore no common quantitative factor was found. No memory factor appeared. The problems which may arise in

preschool assessment due to a limited factor structure have been discussed by Telzrow.

The manual also compared the SBIV to the Wechsler Preschool and Primary Scale of Intelligence (WPPSI). The mean Composite score on the SBIV was 105.3 and the Full Scale score on the WPPSI was 110.3. Thorndike and his associates suggest that this 5 point difference is due to the AVR and Performance IQ differences and the 18 year difference in standardization dates. The combined mean scores on Verbal and Quantitative Reasoning and Short-Term Memory were compared with the Verbal Scale of the WPPSI. The difference between these scores was very small (2.7 points). The largest difference (9.4 points) in mean scores was between the Abstract/Visual Reasoning and the Performance Scale score of the WPPSI, in favor of the Performance IQ. The manual suggests that these differences may be due to several factors. One possible reason for the higher Performance IQ is the difference in standardization between the two tests. Thorndike et al. also suggested that the Matrices test on the SBIV has no counterpart on the WPPSI. As a final comment, the authors of the Binet IV noted that their preschool sample had a greater proportion of girls than boys. Research has shown the girls generally achieve higher scores than boys on 3 out of 5 tests on the Performance scale. This research showed no significant gender differences on the A/VR scale.

Convergent Validity of the SBIV with The WRAT-R, VMI and Pass

WRAT-R

Matuszek (1985) has reviewed current studies on the WRAT and has indicated that the WRAT correlated moderately with various IQ tests, in the range of .40 to .70 for most groups and most tests. Hale (1978) found the WISC-R Verbal IQ to correlate more highly than the Performance IQ with WRAT subtests. Correlations of the Verbal IQ with WRAT tests ranged from .49 to .63. The Performance IQ correlated .26 with WRAT Spelling; .29 with WRAT Reading and .44 with Arithmetic. Using a multiple regression analysis, it was found that the Performance IQ did not significantly predict academic achievement and that the Verbal IQ significantly predicted only Reading and Arithmetic WRAT scores. Fowler (1986) correlated the Verbal Comprehension

(VC) and Perceptual Organization (PO) factors with WRAT subtests for an LD sample. As in Hale's study, the verbal area on the WISC-R was more highly correlated with WRAT scores. Correlations of the VC factor with the WRAT ranged from .61 to .64. The PC correlations with WRAT scores ranged from .16 to .39. Brock (1982) studied a learning disabled sample and correlated the WISC-R subtest scores with WRAT Spelling and Arithmetic scores. Performance subtest scores showed a negligible relationship with both Spelling and Arithmetic tests, with correlation ranges of .01 to .19 and .06 to .31 respectively. The correlations of the WRAT subtests with verbal scores were stronger and ranged from .26 to .47. Smith and Smith (1986) studied the relationship of the WISC-R with the WRAT-R and also studied the WISC-R as a predictor of WRAT-R scores. Their sample included children from a rural school district who were referred for evaluation. The Verbal IQ correlated .41 with Spelling scores, .58 with Reading and .64 with Arithmetic. Correlations of the Performance IQ with Spelling, Reading and Arithmetic scores were .21, .34 and .50 respectively. Multiple regression analysis showed the WISC-R Verbal IQ to be a significant predictor of WRAT-R achievement measures. The Performance scores were not able to predict the WRAT-R scores.

Bossard and Galusha (1979) investigated the relationship of Stanford-Binet LM and the WRAT. Correlations were higher than those described previously for the WISC-R. The SBIQ correlated .77 with Arithmetic, .79 with Reading and .81 with Spelling.

Meloff (1987) recently compared the WRAT-R to the SBIV. All correlations were significantly high. The WRAT-R Reading score correlated .74 with the Composite score. Coefficients of the WRAT-R Reading with the Area scores ranged from .62 to .68. Within subtests, the correlations with the Reading scale ranged from .29 (Memory for Objects) to .66 (Vocabulary).

WRAT-R Spelling correlated highest with the Composite, with a coefficient of .64. Area score correlations were recorded between .53 and .58. The subtest scores showed a more moderate relationship with the WRAT-R Spelling subtests. Subtest Correlations between these two tests ranged from .27 (Memory for Objects) to .58 (Quantitative).

WRAT-R Arithmetic correlations with the SBIV were all high correlations. The relationship between the Composite and Arithmetic scores was recorded at .75. The test area which was most strongly related with Arithmetic was Quantitative Reasoning. The obtained correlation coefficient for these two

scores was .77. Memory for Objects showed the lowest correlation (.38), while Number Series showed the highest correlation (.79).

The use of both these tests for academic evaluation is justifiable based on their moderate relationship. Furthermore, the SBIV appears to be measuring some skills which are related to academic achievement.

VMI

The VMI is often used as a screen for perceptual or motor weaknesses. The relationship of the VMI with other tests has been studied to determine its usefulness as a diagnostic instrument and to provide data for convergent validity studies. A test of visual motor integration is not expected to correlate highly with an intelligence test since they measure different constructs. However, specific visual motor components of intelligence tests such as the WISC-R and SBIV measure similar skills and therefore some relationship between the two tests is expected. Beery (1967) reported correlations between the VMI and mental age on the Primary Mental Abilities test of .59 for first grade children, .37 for fourth grade children and .38 for seventh grade children. Wright and DeMers (1982) found that the VMI correlated more highly with the WISC-R than with WRAT subtests. The VMI correlated .58 with the Verbal IQ, .66 with the Performance IQ and .66 with the Full Scale IQ. The WRAT correlations with the VMI ranged from .33 to .48. However, when a partial correlation was completed and Full Scale IQ scores were partialled out, the WRAT and VMI correlations dropped significantly. Partial correlations of the VMI and WRAT were .13 for Reading; .20 for Spelling and .17 for Arithmetic.

In a convergent validity study of the SBIV, Meloff (1987) found moderate correlations between the Binet and the VMI. There was a low, but significant correlation with the Composite score (.28). The VMI correlated highest with Copying at .61, followed by Matrices with .44. Contrary to expectations, there was no significant correlation with the Abstract/Visual Reasoning score (.24) or with the Pattern Analysis subtest score (.17).

PASS

A summary of the results of recent studies suggested that school achievement is positively associated with level of self-concept (Gurney, 1986). Several studies of the relationship between the PASS and intelligence tests are reported in the PASS manual (Boersma and Chapman, 1985). The relationship of the Full Scale PASS score and the WISC-R IQ scores were investigated with a sample of 73 normally achieving Grade 3 children. Weak and non-significant values were found for PASS correlations with WISC-R Verbal (.12), Performance (.13) and Full Scale (.03) IQ scores. The authors of the PASS suggest that perceptions of school ability are relatively independent of individual intelligence measures. Perceptions of ability may relate to experiences in school, rather than to intelligence.

Meloff (1987) found a modest correlation between the PASS general Ability scale and the Binet A/VR area of .35. Correlations between General Ability and the subtests Copying at .39 and Pattern Analysis at .31 were also reported. No other statistically significant correlations were found between the PASS and Binet Composite, Area scores or individual subtests.

Summary

In this chapter, the development of the Stanford Binet and Wechlser scales was reviewed. Current literature on the Stanford-Binet Fourth Edition was reviewed. Strengths of the new Binet include the ease of administration, straightforward scoring, adaptive testing format, wide age range of administration, good information on reliability and validity, and a large standardization sample (Keith, 1987; Osberg, 1986; Sim, 1987; Vernon, 1987;). The Binet has been criticized for its excorbitant cost (Slate, 1986). Factor analytic data has shown that several subtests do not load on their respective factors and that the hierarchical model of intelligence may not be supported (Slate, 1987; Vernon, 1987). Due to the adaptive format, the inter-examiner reliability of test scores may be reduced (Slate, 1986; Vernon, 1987). Vernon also states that the SBIV is not adequate for work with adults; the content is limited for assessment of preschool children. He also mentions cultural bias as a problem in the new Binet. Consideration of the strengths and weaknesses of the SBIV is relevant to the purposes of this study.

The discriminative validity of the SBIV was one area of investigation of this study. In this chapter, studies on WISC-R pattern analyis of special educational groups were reviewed. General patterns of high and low scores emerged for the learning disabled, mentally handicapped, gifted and normally achieving groups. For example, the learning disabled group tends to have strengths in perceptual and visual motor areas, lower scores in verbal areas and weaknesses in the ACID tests (Arithmetic, Coding, Information and Digit Span). Studies on profile analysis provided background to the SBIV profiles found in this study for special educational groups Recent studies on discriminative validity of the Binet were cited from the manual and the work of Meloff. The manual states that the SBIV is able to differentiate learning disabled from mentally retarded populations and gifted from normal populations. Meloff's study included a description of which SBIV subtests showed significant difference for the four educational groups of learning disabled, mentally retarded, gifted and normal children. The current study provided a replication and expansion of the work of Meloff.

The discriminative ability of the WRAT-R, VMI and PASS were also reviewed. The research suggests that the WRAT-R is able to differentiate among learning disabled, mentally handicapped, gifted and normal groups. The VMI tends to show significant differences between learning disabled and normal groups (Janzen et al., 1983; Koppitz, 1981; Meloff, 1986). Meloff (1987) found no significant differences between the learning disabled and mentally retarded groups or the normal and gifted groups on the VMI. It has been suggested that the PASS scores are lower for learning disabled children (Chapman and Boersma, 1979). Meloff found no significant differences between the LD group and any other educational group. For the MR group, the only significant difference was between MR and normal groups on General Ability and Reading/Spelling Scales. Meloff found no significant differences between the gifted group and any other educational group. The discriminative ability of these tests was investigated in this study and compared to the current literature.

Studies pertaining to the convergent validity of the SBIV were reviewed in this study. Results of studies in the Manual indicate that the SBIV is strongly related to the WISC-R. The relationship between the WRAT-R and the SBIV is moderately strong (Meloff, 1987). Previous studies show moderate correlations between the WRAT and intelligence tests (Bossard and Galusha,

1979; Brock, 1982; Fowler, 1986; Hale, 1978; Matuszek, 1985; Smith and Smith, 1986). The VMI shows moderate relationships with intelligence measures (Beery, 1967; Wright and DeMers, 1982). Meloff found low to moderate correlations between the Binet and the VMI. Previous studies of the PASS have shown it to be weakly related to intelligence measures. Many of the obtained correlations are nonsignificant (Boersma and Chapman, 1985). Meloff found a very low correlation between the PASS general Ability scale and the Binet A/VR area of .35. Correlations between General Ability and the subtests Copying at .39 and Pattern Analysis at .31 were reported. No other significant correlations were obtained.

The use of the Binet with preschoolers was also reviewed in this chapter. Telzrow (1986) and Schakel (1986) discuss the strengths and weaknesses of the SBIV with this group. The manual provides factor analytic data and a comparative study with the WPPSI. These studies were reviewed for comparison with results obtained in this data.

Current studies on the SBIV show it to have good discriminative and convergent validity. The purpose of this study was to confirm these findings and provide further information on the validity of the Binet IV.

CHAPTER THREE

Research Design and Procedures

Procedure

Data Collection

Data for this thesis were collected by graduate students who were enrolled in the Fall/Winter and Spring/Summer sessions of the Educational Psychology 524 individual assessment course at the University of Alberta.

Subjects in the study from the Fall/Winter session were administered the Stanford-Binet Intelligence Scale: Fourth Edition (SBIV), the Wechsler Intelligence Scale for Children-Revised (WISC-R) and additional tests that comprised the test battery. Administrations of the SBIV and the WISC-R were counterbalanced for order with an interval of 4 to 8 weeks between the tests. Each graduate student was required to give both tests to all assigned clients. In cases where the age norms were not appropriate, the WISC-R was not given. This excluded children below the age of 6 years and above the age of 17 years. Some clients did not wish to have both the SBIV and the WISC-R administered. In these cases, only the SBIV and additional tests were given. Subjects in the study who were seen in the Spring/Summer Session were administered only the Stanford-Binet Intelligence Scale: Fourth Edition and additional tests. The graduate students were competently trained in the administration, scoring and interpretation of all the tests in the respective batteries and as a standard procedure of the Educational Psychology 524 course were supervised in all aspects of their work by a trained senior clinician. The data for this study were collected by reviewing the subjects' test reports. These reports are held within confidential files in the clinic and are only available for students registered in the appropriate courses. Files are locked after regular clinic hours. Test report information was coded onto data sheets and entered into the university computer system. A statistical analysis of the data was completed using programs from the SPSSX package available on the Amdahl computer at the University of Alberta.

Education Clinic

The subjects for the study were clients from the Education Clinic at the University of Alberta who were either self-referred or referred by parents, teachers, administrators or health professionals. Services at the clinic are free of charge. Student clinicians perform assessments and provide counselling services. The clients utilize the services of the clinic with the understanding that information gathered in a session may be used for teaching or research purposes. There are no age limitations for clients who use the clinic services. The public is provided with information about the clinic through a radio program which is hosted by university professors from the Educational Psychology department, through the school system and through other public sources. Clinicians are available to answer questions about clinic services during regular weekly office hours.

Referrals for the clinic are received over the phone or are taken directly at the clinic. Clients are seen in order based on the date when the referral was accepted. There is usually a short waiting time before the clients are seen by student clinicians. The lapse time may be from a few weeks to a few months. After the assessments are completed, the clients are provided with a debriefing session at which time the tests which were given are described, interpretations of the results are given and recommendations are made. In addition to clients coming to the Education Clinic for services, graduate students go outside the agency to assess clients in homes, schools, hospitals, or other institutions. Assessments are completed both within Edmonton and in surrounding cities and outlying communities.

Sample

The sample consisted of 326 subjects. There were 214 subjects who had been seen by student clinicians attending the Fall/Winter session of the Educational Psychology 524 assessment course. From the Spring/Summer Session, 112 subjects were added to the sample. Table 3.1 reports the tests administered and what portion of the sample was given each test.

TABLE 3.1

Distribution of the Sample by Tests Given

TESTS	N
SBIV and WISC-R	173
SBIV only	326
SBIV and WRAT-R	224
SBIV and PASS	120
SBIV and VMI	134

Appendix A reports the distribution of clients in terms of demographics. The subjects ranged in age from 2 years, 2 months to 27 years, 6 months (mean age = 9.6 years). In this study there were 187 male subjects (57.4%) and 139 female subjects (42.6%).

Five ethnic groups were identified in this sample. The predominant group was Caucasian and made up 83.7% of this sample. The other ethnic groups which were identified included Native, Asian, Hispanic, and other .

Approximately 73 percent of the population in this study came from cities. Subjects from small towns and rural areas totalled close to 27% of the sample.

Parental occupations were listed within seven categories. Six of these were the same categories used in the Technical Manual (Thorndike et al., 1986b). One extra category, "unemployed", was added to this study. The Managerial/Professional Category was represented by the largest number of parents (27.3%) followed by Technical / Sales and Service Occupations which contributed about 25% each. Based on the standardization sample presented in the Technical Manual, the Managerial/Professional category of this sample appears to be slightly over-represented. The Service category also showed an over-representation based on the figures provided in the manual. However, the Canadian sample appears to be more similar to weighted sample of the manual than to the actual standardization sample.

In the Parental Education category, 1/4 of this sample had 4 or more years of post-secondary education. Parents with one to three years of college education represented approximately 20% of the sample in this study. High School graduates also represented close to 20% of the sample. Only 7% of

this sample were tabulated as having less than a high school education. This may be an under-representation of this category. It is much less than the weighted American Sample percent described in the Technical Manual. However, all of these figures for this sample must be noted as being estimates only, since some data was missing. The educational level of 27.6% of the sample was unknown.

With regard to parental marital status, the majority of the sample (84.7%) were married at the time of testing.

Referral reasons for each subject were taken from the "Report Forwarding Form" (Appendix B) which is used at the Psychological Testing Center at the University of Alberta. Referral reasons listed on the form which were included in this study were the following: readiness for kindergarten, pre-school assessment, learning disabilities (child), learning disabilities (adult), dyslexia (child or adult), school learning problems (general), behavior problems, developmental delay/maturation, social/emotional problems (general), self-referral/general interest, career choices/planning, reading/language problems, giftedness/enrichment, Parental Interest (in child's ability), special class or program placement, special school placement and mental retardation. For the purpose of this study, these categories were regrouped into five subgroups (see Table 3.2). A description of the referral reasons relevant to each category follows.

1. Learning Difficulties (Learning disabled) included the following referral reasons: learning disabilities (child), learning disabilities (adult), dyslexia (child or adult), school learning problems (general), developmental delay / maturation, and reading/language problems. A description of the operational definition of learning disabilities which was used in this study is included below.

In the United States, it is the requirement of Public Law 94-142 that individuals with any handicapping conditions be identified, assessed and provided with educational opportunities which maximize their strengths. Included in P. L. 94-142 is a definition of learning disabilities which is commonly used by researchers, educational psychologists and teachers:

> "Specific learning disability" means a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which may manifest itself in an imperfect ability to listen, think, speak, read, write, spell or to do mathematical calculations. The term includes such conditions as perceptual handicaps, brain injury, minimal brain dysfunction, dyslexia and developmental aphasia.
> The term does not include children who have learning problems which are primarily the result of visual, hearing, or motor handicaps, of mental retardation, of emotional disturbance, or of environmental cultural or economic disadvantage.
> (Federal Register, December 29, 1977 in Sattler).

This definition has been criticized for its limitation of the disability to children. It has been suggested that a definition of learning disabilities should also include adults and youth. Furthermore, exclusion of visual, hearing or motor handicaps does not allow for cases in which there may be physical handicaps in addition to a learning disability. These two concerns were raised by the National Joint Committee on Learning Disabilities (NJCLD, 1987). The previous definition was used for the classification purposes of this study. However, the criticisms of the NJCLD were used to make slight amendments, that is, adults and youth were also included in the learning disabled study. Individuals with physical handicaps secondary to a learning disablility were included in the study.

2. Mental Retardation included the following referral reasons: mental retardation, special class or program placement and special school placement. Many subjects in this group were referred from the Edmonton Catholic school system and were enrolled in programs fof the Educably Mentally Handicapped (EMH) at the time of testing. Therefore, this group was referred to as the mentally handicapped (MH) group in this study. The operational definition used for this study is the one used by the American Association on Mental Deficiency (AAMD) (Grossman, 1973, in Sattler, 1982)

> Mental Retardation refers to significantly subaverage general intellectual functioning existing concurrently with deficits in adaptive behavior, and manifested during the developmental period (p. 11).

3. Gifted included the following referral reasons: giftedness/enrichment. In Alberta, the following definition is used to describe gifted children. This definition is the one which was used for the purposes of this study.

> Gifted and talented children are those who, by virtue of outstanding abilities, are capable of high performance. These are children who require differentiated educational programs and/or services beyond those normally provided by the regular school program in order to realize their contribution to self and society (Wilde and Sillito, 1986, p.7).

To qualify for inclusion in an academic challenge program in the Strathcona County school system a student must demonstrate some or all of the following qualities:

1. Superior intellectual ability as obtained on an individually administered intelligence test (IQ of 125).
2. Superior academic achievement. Preferable one and one-half to two years above grade level in total reading score as measured by the Canadian Achievement Test.
3. Adequate social and emotional maturity
4. Ability to work independently with little rigid discipline
5. Enjoyment of intellectual challenges: performs well on difficult tasks
6. Maintenance of a satisfactory level of achievement in the regular classroom (Wilde and Sillito, 1986, pp. 9-10).

4. Parental Interest (Normal) included the following referral reasons: parental interest (in child's ability), behavior problems, social/emotional problems (general), self-referral- general interest and career choices/planning. This category functioned as a normal control group within the study. All subjects were within the average IQ range (89 - 110) or the above average range (111-120) as defined by the classification guide of the Stanford-Binet Fourth Edition. There were no academic delays or deficits within this group of subjects.

5. Pre-School included the following referral reasons: readiness for kindergarten (4-5), pre-school assessment.

Table 3.2
Reason for Referral by Subgroup

	Subgroup	Number Percent
Learning Difficulties	93	28.5
Mentally Handicapped	46	14.1
Gifted	59	18.1
Parental Interest	68	20.9
Pre-School	60	18.4
TOTAL	326	100.0

Tests Administered

Stanford-Binet Intelligence Scale: Fourth Edition (SBIV); Wechsler Intelligence Scale for Children-Revised (WISC-R); Wide Range Achievement Test-Revised (WRAT-R); Developmental Test of Visual Motor Integration (VMI)); and the Perception of Ability Scale for Student's (PASS). All scores cited in the study are standard scores except for in relationship to the PASS where only raw scores were available.

Stanford-Binet Intelligence Scale: Fourth Edition

Theoretical Model

The authors used a three-level hierarchical model to construct the SBIV. A general reasoning factor or "g" is at the top. Crystallized abilities, fluid-analytic abilities and short-term memory form the middle level. The third level is made up of more specific factors which include verbal reasoning, quantitative reasoning and abstract/visual reasoning. Further details are available in the Guide. Specific information in relation to format of the Binet IV, changes from the Stanford-Binet Form L-M and administration guidelines are found in the Guide and the Technical Manual.

Subtest Descriptions

Janzen (1986) has outlined the characteristics of each SBIV subtest. An understanding of each subtest is very relevant to the purposes of this study.

1. Vocabulary: measures expressive language and the ability to define words. Conceptual understanding of both concrete and abstract words is required. Two levels of the Vocabulary subtest are included in the Binet IV for different age levels.

   a. Picture Vocabulary: the description of concrete objects is required. Objects must be distinguished from their function. Perception of whole objects and recognition of categories is tapped.

   b. Vocabulary: Janzen describes this as the most important subtest on the Stanford-Binet. Verbal fluency and abstract thinking are measured. The relationship between a symbol and its meaning is demonstrated. Development of temporal concepts is measured. Long-term memory for linguistic symbols is necessary.

2. Bead Memory: measures the ability to attend and to retain information. Both motivation and attention are required. Visual short-term memory and understanding of spatial relations as well as spatial orientation are measured. Osberg (1986) describes this test as tapping simultaneous memory function.

3. Quantitative: measures counting ability for one-to-one relationship of numbers to objects. Processes involved in this subtest include visual-motor abilities, short-term visual memory, and eye-hand coordination. Sequencing of information which is briefly presented is necessary. To solve numerical problems sequential strategies and short-term memory processes must be evoked. Mental visualization of the problems, language and visualization fo linguistic materials are all involved. Abstract thinking, understanding of analogies and inductive reasoning are all necessary.

4. Memory for Sentences: requires good auditory reception and attentional skills. Short-term memory for language structured stimuli is measured. Auditory semantic sequencing, good word knowledge and well developed articulation are gauged.

5. Pattern Analysis: measures visual-motor skills, spatial orientation, eye-hand coordination, non-verbal and visual-spatial reasoning. Abilities in the area of form perception, part-whole relationships, gestalt perception and analysis of angular patterns are estimated. Problem solving strategies and a trial and error approach are demonstrated during completion of the task.

6. Comprehension: requires integration of verbal and non-verbal information to form a common sense solution. Language skills and verbal fluency are tapped as social confidence and judgment are expressed. Analysis of a question requires adequate receptive language skills and social insight. There is a need for alertness to the environment, visualization and the capacity for mental imagery. Long-term memory for language and capacity to shift a mental set are required.

7. Absurdities: Visual awareness, reasoning, long term memory for visual objects and understanding of spatial relations and orientation are necessary. Social awareness and understanding of social and physical conventions are tapped. The association between visual stimuli and verbal reasoning is estimated.

8. Memory for Digits

   a. Digits Forward: measures auditory short-term memory, rote learning, aural sequencing skills and attentional skills.

   b. Digits Backwards: measures the same skills required in Digits Forward but also requires mental reordering of stimuli without losing the memory trace.

9. Copying: This subtest measures the client's ability to manipulate concrete material, integration of visual and motor skills, and eye-hand coordination. Perception of geometric forms, gestalt perception, understanding of spatial relations and figure ground perception are all required. Trial and error problem solving is also measured.

10. Memory for Objects: measures short-term visual memory, and sequencing skills. Visual reception, the recognition of objects and expressive language are tapped. It involves a process of recognition memory, not simply recall.

11. Matrices: requires non-verbal, visual-spatial reasoning. Observation of details, visual acuity and visual perception are necessary. It measures abstract visual reasoning, attention to sequence, form perception and

understanding of part-whole relationships. Inductive reasoning skills and the capacity to shift mental set are required.

12. Number Series: measures analytical thinking, mathematical concept formation, and understanding of a mathematical rule or principle. Sequencing skills, the knowledge of numerical relations, fractions and multiplication are necessary.

13. Paper Folding and Cutting: measures inductive visual reasoning. It requires analytic thinking, sequencing of visual data and attention to detail. Understanding of spatial relationships, visual closure and gestalt perception are tapped.

14. Verbal Relations: measures verbal reasoning, associative thinking and concept formation. It requires good long-term memory for semantics. It is necessary to have the capacity to shift mental sets, be aware of word opposites and have a good word knowledge.

15. Equation Building: requires knowledge of basic arithmetic operations and functions: long-term memory for math facts and operations. Must be able to mentally reorganize numerical data, visualize alternative sequences and shift mental sets. Abstract/visual reasoning is also required.

Reliability

Reliability for the SBIV was reported on the basis of internal consistency and test-retest reliability. The information on internal consistency was based on the Kuder-Richardson Formula 20 (KR-20). The most reliable score was the Composite SAS, with obtained reliability indices ranging from .95 to .97. The next most reliable scores were the Area scores. For the Verbal Reasoning Area, the reported reliabilities ranged from .90 to .97. Short-Term Memory ranged from .86 to .95 and Abstract/Visual Reasoning from .85 to .96. The Quantitative Reasoning Area obtained the lowest reliability indices, from .81 to .97. The authors of the manual noted that the reliability of an Area varied according to the number of tests that were included in it, which was also dependent on the age of the client and adaptive testing patterns. The reliabilities for specific subtests were slightly lower than for the Areas and Composite. These reliabilities were typically in the .80s and .90s. The one exception was Memory for Objects, for which reliabilities ranged from .66 to .78. This is a short test composed of only 14 items. Tables 5.1 to 5.3 of the

Technical Manual (Thorndike et al, 1986b) show further information on internal consistency. There was a general tendency for reliabilities to be higher for the older age groups.

Test-retest reliability data was obtained from retesting 112 children. The sample consisted of two groups - a group of 57 children first tested at 5 years of age and a group of 55 children first tested at 8 years of age. The test-retest reliabilities for all scores were lower than the internal consistency reliabilities for both groups. This difference may be due to the small sample size and the restricted range of the test-retest groups. The test-retest correlations for the Composite Score were .90 for age 8 and .91 for age 5. Area score reliabilities ranged from .71 to .88 for age 5. For the 8-year old group, the Area score reliabilities ranged from .51 to .81. For both age groups, the Quantitative Reasoning Area showed the lowest reliability. Subtest reliabilities ranged from .56 (Bead Memory) to .78 (Memory for Sentences) for age 5; from .28 (Quantitative) to .86 (Comprehension) for age 8. Consult tables 5.4 and 5.5 of the Technical Manual for further information.

Validity

Construct validity is reported in the SBIV manual based on three specific areas:

1. The internal structure of the test as evidenced by the correlations among the tests and as clarified by factor analytic procedures.
2. The correlation of the test with scores on other tests deemed to be measures of the same or a similar construct; and
3. the performance of groups identified by indices other than the Fourth Edition as presumably high or low on the construct(s) the test is designed to measure.

(Thorndike et al., 1986b, p.52.).

Intercorrelation coefficients were calculated between all the tests administered to each of the 17 age groups in the standardization sample. Medial intercorrelations are reported in Table 6.1 of the Technical Manual. Complete information for all age groups is reported in Appendix B of the manual. The intercorrelation coefficients of the Composite score with the area scores

ranged from .86 to .90. Subtest intercorrelations with the Composite score ranged from .60 to (Memory for Objects) to .81 (Vocabulary).

A variant of confirmatory factor analysis was carried out on the median correlations. All tests received substantial loadings on "g" or the General Ability factor. Tests with the highest "g" loadings were Number Series (.79), Quantitative (.78), Vocabulary (.76) and Matrices (.75). Tests with relatively low loadings were Memory for Objects (.51), Memory for Digits (.58) and Copying (.60). Each of the four Areas were confirmed as valid factors. The group factor loadings were all lower than the general factor loadings. The specificity of each test was calculated using KR-20 values. All the tests measured some specific variance. Copying had the highest value at .69. Additional factor analyses were performed for three age groups - 2 through 6, 7 through 11, and 12 through 18-23. Details are contained in the manual.

Correlations with other tests were reported in the manual for the Stanford-Binet L-M, WISC-R, WAIS-R, WPPSI, and Kaufman Assessment Battery for Children (K-ABC). Details of comparisons between the WISC-R and the SBIV were given in chapter two of this study. Details of other comparative studies are available in the manual.

Wechsler Intelligence Scale for Children-Revised

Description

The WISC-R is divided into two main areas which include verbal and performance (visual-motor) tasks. There are 12 subtests in total, two of which are supplementary tasks and are not usually included in the calculation of the IQ score. The Verbal and Performance areas include six subtests each. All subtests are given to all age levels. The WISC-R is intended for use with children from 6-0 to 16-11 years. The subtest scores have a mean of 10 and a standard deviation of 3. The subtest scores are combined in each area to form a Verbal IQ and Performance IQ respectively. The combination of ten subtest scaled scores is used to obtain the Full Scale IQ. The IQ scales have a mean of 100 and a standard deviation of 15. Approximately 50 to 75 minutes are required to administer the regular battery of 10 tests. Details regarding the standardization sample, administration procedures and format are available in the WISC-R manual (Wechsler, 1974).

Reliability

Reliability was based on internal consistency and test-retest measures of the WISC-R. For every test except Digit Span and Coding, reliability coefficients were obtained by the split-half technique and corrected with the Spearman-Brown formula. Digit Span and Coding reliabilities were based on test-retest procedures. The reliability coefficients for the Verbal and Performance IQs were based on a composite of the five subtests of each area; the reliability of the Full Scale IQ was based on ten tests. The Verbal, Performance and Full Scale IQs had high reliabilities across the entire age range, the average coefficients being .94, .90 and .96 respectively. The average coefficients ranged from .77 to .86 for the Verbal tests and from .70 to .85 for the Performance tests.

Test-retest reliability coefficients for the Verbal, Performance and Full Scale IQs were .93, .90 and .95 respectively. The average stability coefficients ranged from .77 to .88 for the Verbal tests and from .65 to .81 for the Performance tests.

Validity

Validity information was based on intercorrelation of the tests and on correlations with other measures of intelligence. The intercorrelations of the twelve WISC-R tests for each of the eleven age groups were obtained. The correlations of each test with the Verbal score, Performance score and Full Scale score were also obtained. The Verbal score was also correlated with the Performance IQ. The average intercorrelations of subtests with the Full Scale IQ ranged from .38 to .74. Average Performance subtests correlations with the Performance IQ ranged from .33 to .68. The Verbal IQ correlated with verbal subtests from a range of .45 to .78.

Special studies were undertaken to determine the relationship between the WISC-R and three other individually-administered intelligence tests: the WPPSI, the WAIS and the Stanford-Binet Intelligence Scale, Form L-M. The correlation between the WISC-R Full Scale IQ and the WPPSI Full Scale IQ was .82. The WISC-R Verbal IQ correlated .80 with the WPPSI Verbal IQ. The WISC-R Performance IQ correlated .80 with the WPPSI Performance IQ. The

correlation between the WISC-R and WAIS Full Scale IQ was .95, the correlation between the two Verbal IQs was .96 and the correlation between the two Performance IQs was .83. The average coefficients of correlation of the WISC-R Verbal, Performance and Full Scale IQs with the Stanford-Binet IQ were .71, .60 and .73 respectively.

Wide Range Achievement Test-Revised

Description

Joseph J. Jastak was responsible for the development of the original Wide Range Achievement Test standardized in 1936. The WRAT was developed as a quick-screening instrument to evaluate achievements in the areas of reading (word recognition, naming letters, and pronouncing words), spelling (copying symbols, writing one's name and spelling words from an oral dictation), and arithmetic (counting, reading numbers symbolically, and solving oral and written computations).

Five revisions have been done over the past 50 years. The most recent revision was completed in 1984 by Sarah Jastak and Gary Wilkinson. A restandardization was completed at this time, using six thousand subjects. The tasks of the WRAT-R are similar to those of the original WRAT. The WRAT-R has two forms: Level I for subjects aged 5-0 to 11-11, and Level II for subjects aged 12-0 to 74-11. The test is aged normed which means that each individual's scores can be compared to the scores of other members of their own age group. Each level takes approximately twenty to thirty minutes to administer in an individual session. The WRAT-R generates four kinds of scores: raw, grade equivalents, standard scores and percentiles. Standard scores are deviation quotients with a mean of 100 and a standard deviation of 15. These scores can be compared directly with IQ scores.

Reliability and Validity

Test-retest coefficients for the WRAT-R Level I Reading, Spelling and Arithmetic tests were reported to be .96, .97 and .94 respectively. For Level II, the stability coefficients were slightly lower. For Reading, Spelling and Arithmetic, the coefficients were .90, .89 and .79 respectively (Jastak and Wilkinson, 1984, p.64.)

Convergent validity studies cited in the manual compared the WRAT with the Peabody Individual Achievement Test, the California Achievement Test and the Stanford Achievement Test. Correlations of these measures ranged from the high .60's to the .80's. These tests of validity were cited for the WRAT rather than the WRAT-R as information for the WRAT-R in this area was not available at the time of publication.

The Developmental Test of Visual Motor Integration

Description

This test was originally developed by Beery in 1967 and revised in 1982. The VMI is a perceptual motor ability test for children from ages 2 to 14 years. It contains 24 geometric forms of increasing complexity which are copied by the child. Each design is scored on a pass or fail basis. An age equivalent is estimated from the total raw score. A scaled score is also obtained which has a mean of 10 and a standard deviation of 3, thus making comparisons with the WISC-R possible. The test can be administered in individual or group sessions in about 15 minutes. The current 1982 norms are based on a sample of 3,090 children.

Reliability

Test-retest reliabilities of the VMI are reported in the manual (Beery, 1982) to range from .63 (seven month period) to .92 (two week period). The median reliability cited in the manual is .81. These reliabilities are based on studies with normal subjects. Split half reliabilities which are cited in the manual ranged from .66 to .93 with a median of .79.

Validity

Correlations between VMI scores and chronological age are reported in the manual to be .89 for the two to 15 year age range. Studies cited in the manual report correlations between the VMI and Bender test to have ranged from .41 to .82. Correlations between the VMI and readiness tests have averaged about .50. Correlations with reading and other academic

have tended to be higher for the primary grades than for the upper grades. Correlations between the VMI and mental age on the Primary Mental Abilities test were .59 for the first grade, .37 for the fourth grade and .38 for the seventh grade.

Perception of Ability Scale for Students (PASS)

Description

This test assesses cognitive self concept. It can be used on school aged children from grades three to six. A brief questionnaire contains items related to general feelings about school, and feelings related to abilities in mathematics, spelling, reading, printing and neatness. Responses to the items are given by circling "yes" or "no" beside each question. Seven scales make up the PASS and include (1) general ability (2) arithmetic (3) reading/spelling (4) school satisfaction (5) penmanship and neatness (6) confidence and (7) Full Scale Score. Raw scores for each of the scales are obtained from the protocol.

Reliability

Estimates of internal consistency were obtained by using Cronbach's alpha. This analysis was based on data from the normative sample of 642 children in grades 3 to 6. The Full Scale alpha was .915. Arithmetic, Reading/Spelling and Penmanship/Neatness alpha estimates were between .822 and .855. For General Ability the alpha was .785, School Satisfaction was .741 and Confidence was .686.

Test-retest reliability data were collected after a 4 to 6 week interval, using a sample size of 603 children. The stability coefficient for the Full scale score was .834. Subscale values ranged from .714 to .824.

Validity

The authors of the PASS discussed validity studies using the following measures:

1. School achievement based on report cards. Some specific PASS subscales had moderate correlations with report card grades. These correlations ranged from .27 to .37.
2. Standardized tests (Canadian Test of Basic Skills) Correlations ranged from .12 to .44 with a number of negative correlations reported.
3. Intelligence measures. The PASS had very low nonsignificant relationships with IQ scores. The authors suggest that perception of school ability may be relatively independent of individual and group intelligence measures.

Analysis of the Data

Missing Data

The SBIV has an adaptive format which means that not every subject is required to complete all 15 subtests. Furthermore, not all tests were administered to all of the subjects in this study. Table 3.3 includes the list of the variables and the percentage of data missing. For the WISC-R subtests, the percentage of missing data is based on a partial sample which includes the 214 subjects seen the the Fall/Winter Session of the Educational Psychology assessment course. Treatment of the missing data varied with each analysis and is discussed separately with presentation of each analysis.

Description of Statistical Analysis

Frequencies and percentages were generated for the total sample in terms of demographic variables of age, gender, ethnicity, locale, parental occupation, parental education and parental marital status. These variables are reported in Appendix A, Tables I to VII.

The variable "Reason for Referral" (Table 3.1) was reported in terms of frequencies and percentages for the five subgroups of learning difficulties (learning disabled), mentally handicapped, gifted, parental interest (normal) and pre-school subjects.

Summary statistics for the entire sample are shown in chapter four. Table 4.1 contains means and standard deviations for the SBIV and the WISC-R. Table 4.2 includes the means and standard deviations for the WRAT-R, VMI and PASS variables. Subtest mean scores for each group were

evaluated individually to obtain group profiles of high and low scores. Table 4.3 contains means and standard deviations for the four groups. Table 4.4 shows the high and low scores for the learning disabled, mentally handicapped, normal and gifted groups.

A linear regression analysis was carried out using the SBIV subtests as the independent variable and the groups (MH, LD, N and G) as the dependent variable. A stepwise procedure was utilized, followed by a forced entry method. In this manner, it could be determined which subtests were the best predictors of educational category and what percentage of variance in the educational category the SBIV subtests predicted. Results of the multiple regression are shown in Table 4.8. For the regression analysis, missing data was subjected to the "pairwise missing-value treatment" (SPSS Inc., 1983, p. 606). Using this method, each correlation coefficient is calculated using cases with complete data for the two variables, regardless of whether the cases have missing data on any other variable.

A Linear Discriminant Function was also used to determine which combination of SBIV subtests would best predict membership in each of the four groups. Results are shown in tables 4.9 to 4.12. For the linear discriminant function, cases with missing values were not used during analysis. During classification, the means of variables were substituted for missing values and cases containing missing values were classified (SPSS Inc., 1983, p.641). Without using this method, the number of cases available for analysis would be reduced to a miniscule size.

Pearson Product Moment Correlations were calculated to determine the relationships between the SBIV and the WISC-R for subtest scores, and Area and IQ scores. Pearson Product Moment Correlations were also calculated to determine the Stanford-Binet's relationship with the WRAT-R, VMI and PASS. The results are indicated in tables 4.13 and 4.14. Cases with missing values were deleted on an "analysis-by-analysis" basis. That is, a case missing for either part of the pair is not used in the correlational analyses. This results in a matrix which is based on a varying number of cases (SPSS Inc., 1983, p. 584).

A oneway analysis of variance (anova) was used to determine whether differences existed between the mean scores attained by learning disabled, mentally handicapped, gifted and normal subjects. Anovas were used to determine differences on the WRAT-R, VMI and the PASS. The anova results

are shown in Appendix C, Tables IX to XXXV. The Scheffe test, a comparison test, was used in this study to see which population means differed from one another. Results of the Scheffe test are presented in Tables 4.5 and 4.6. Missing values were deleted on an "analysis-by-analysis basis."

The subtest scores on the SBIV were placed in ordinal position from highest to lowest for Pre-school children and evaluated as a "typical profile." Procedures were similar to the clinical evaluation of the special educational scores. Composite and area scores were compared to results in the manual.

Table 3.3
Percentage of Missing Data

Variable	Percentage of Data Missing
WISC-R Subtests	17.7
WRAT-R Subtests	30.4
VMI	58.0
PASS	62.8
SBIV Subtests	
Comprehension	1.8
Absurdities	10.4
Verbal Relations	88.0
Pattern Analysis	.3
Copying	22.4
Matrices	47.0
Paper Folding and Cutting	86.2
Quantitative	3.1
Number Series	5.4
Equation Building	91.7
Bead Memory	.9
Memory for Sentences	2.2
Memory for Digits	31.6
Memory for Objects	42.6
Quantitative SAS	1.5
A/VR SAS	.3

Research Questions and Hypotheses

The following questions and hypotheses formed the basis of the study:

Research Question 1

Will learning disabled, mentally handicapped, gifted and parental interest (normal) groups demonstrate patterns of Binet subtest scores which have distinctive high and low subtest scores for each respective group?

Hypothesis 1

Scores from the SBIV, WRAT-R, VMI and PASS will be significantly different for learning disabled, mentally handicapped, gifted and parental interest (normal) groups.

Hypothesis 2

Scores from the SBIV subscales and area scores will be predictive of membership in learning disabled, mentally handicapped, gifted and normal educational groups.

Hypothesis 3

The Verbal Reasoning SAS, Quantitative Reasoning SAS and Short-Term Memory SAS will correlate more highly with the Verbal IQ and Full Scale IQ of the WISC-R than with the Performance IQ.

Hypothesis 4

The Abstract/Visual Reasoning SAS will correlate more highly with scores on the Performance Scale than with scores on the Verbal Scale.

Hypothesis 5

The Composite score on the Fourth Edition will correlate highly with all three IQ scores on the WISC-R.

Hypothesis 6

The SBIV Vocabulary subtest will correlate highly with the WISC-R Vocabulary subtest.

Hypothesis 7

The Comprehension subtests from the Binet and WISC-R will be highly correlated with each other.

Hypothesis 8

The Memory for Digits subtest from the Stanford-Binet will be highly correlated with the Digit Span subtest on the WISC-R.

Hypothesis 9

The Binet subtest, Pattern Analysis, will be highly correlated with the Block Design subtest on the WISC-R.

Hypothesis 10

The Quantitative Reasoning Area, Number Series subtest and the Quantitative subtest will correlate significantly with the WISC-R Arithmetic subtest.

Hypothesis 11

The WRAT-R Reading and Spelling subtest will correlate significantly with the SBIV Composite IQ, Verbal Reasoning SAS and subtests from the Verbal Reasoning Area.

Hypothesis 12

The WRAT-R Arithmetic test will correlate significantly with the Binet Composite IQ, Quantitative Reasoning SAS, and the Quantitative and Number Series subtests.

Hypothesis 13

The PASS subscale scores will correlate significantly with the SBIV Composite IQ and certain subtests on the SBIV.

Hypothesis 14

The VMI score will correlate significantly with the Composite IQ on the Binet, the A/VR SAS and the Pattern Analysis, Copying and Matrices subtests.

Research Question 2

Will the Pre-school group demonstrate a distinctive profile of high and low scores? How will the results of this study compare to the results of the Technical Manual?

CHAPTER FOUR
RESULTS

The purpose of this study was to investigate the validity of the Stanford - Binet Fourth Edition. The discriminative ability of the SBIV was one primary area of investigation. The manual claimed that this test could discriminate between learning disabled and mentally retarded groups and between gifted and normal groups. The potential use of the Binet scores to make educational decisions required that these claims be confirmed and that further information be gathered. Specifically, the ability of this test to differentiate among mentally handicapped, learning disabled, gifted and normal subjects was a major part of this study. Profiles of high and low scores for each education group were studied in the hopes of finding typical patterns such as have been demonstrated for the WISC-R test. The use of SBIV subtests and area scores to predict membership in a category was also investigated. The efficacy of other supplementary tests, the WRAT-R, VMI and PASS, for making educational decisions was also questioned. With information regarding the discriminative validity of the SBIV, educational psychologists can arrive at more accurate classification decisions. Convergent validity was also an important division of the research. The relationship between the WISC-R and SBIV scores and the respective relationships of the WRAT-R, VMI and PASS with the SBIV were examined.

Descriptive Sample Statistics

Total sample means and standard deviations of the Binet IV and WISC-R variables are reported in Table 4.1. Table 4.2 lists the variable means and standard deviations for the WRAT-R, VMI and PASS. These tables also include variable names and abbreviations for each variable. Further tables and figures in the chapter use abbreviations only. For the educational groups, the following abbreviations are used in the tables in this chapter: MH = Mentally Handicapped, LD =Learning Disabled, N = Normal and G = Gifted.

TABLE 4.1

Total Sample Frequencies, Means and Standard Deviations for Stanford-Binet Fourth Edition and WISC-R Variables

ABBREVIATION	VARIABLE NAME	N	MEAN	STD. DEV
SV	SBIV Vocabulary	326	51.291	9.123
SC	SBIV Comprehension	326	51.684	9.524
AB	SBIV Absurdities	293	51.737	8.839
PT	SBIV Pattern Analysis	325	51.634	9.463
CP	SBIV Copying	254	47.024	9.331
MT	SBIV Matrices	177	50.757	8.837
QU	SBIV Quantitative	321	50.302	9.317
NS	SBIV Number Series	151	51.291	8.514
BM	SBIV Bead Memory	323	49.000	10.461
MS	SBIV Mem. for Sentences	319	48.483	10.173
MD	SBIV Mem. for Digits	223	49.991	8.691
MO	SBIV Mem. for Objects	200	50.215	8.340
VR	SBIV Verbal Reasoning IQ	326	103.696	19.033
AVR	SBIV Abstract/Vis. R. IQ.	325	100.982	19.597
QR	SBIV Quant. Reas. IQ.	321	102.156	18.923
STM	SBIV Short-Term Mem. IQ	326	98.442	20.630
SBIQ	SBIV CompositeIQ	326	101.644	20.488
IN	WISC-R Information	173	8.885	4.059
SI	WISC-R Similarities	173	10.410	4.735
AR	WISC-R Arithmetic	173	8.538	3.803
WV	WISC-R Vocabulary	173	9.474	4.443
WC	WISC-R Comprehension	173	10.231	3.992
DS	WISC-R Digit Span	170	8.629	3.360
PC	WISC-R Picture Comp.	173	10.728	3.625
BD	WISC-R Block Design	173	10.416	4.277
OA	WISC-R Object Assembly	173	9.821	3.845
CO	WISC-R Coding	169	8.438	4.333
MA	WISC-R Mazes	160	11.000	3.925
VIQ	WISC-R Verbal IQ	173	96.913	23.473
PIQ	WISC-R Performance IQ	173	101.110	21.602
FSIQ	WISC-R Full Scale IQ	173	99.151	23.500

TABLE 4.2
Total Sample Frequencies, Means and Standard Deviations
for WRAT-R, VMI and PASS Variables

ABBREVIATION	VARIABLE NAME	N	MEANS	STD.DEV
WR	WRAT-R Reading	224	96.027	20.335
WS	WRAT-R Spelling	223	92.646	19.343
WA	WRAT-R Arithmetic	223	93.767	20.709
PFS	PASS Full Scale	120	43.975	12.326
GA	PASS General Ability	119	6.748	3.092
AT	PASS Arithmetic	119	8.445	2.866
SS	PASS Sch. Satisfaction	119	7.681	2.858
RS	PASS Reading/Spelling	119	7.773	3.446
PN	PASS Penmanship/Neatness	119	8.521	2.911
CON	PASS Confidence	119	4.739	2.309
VMI	Visual Motor Integration	134	8.813	3.447

The Discriminative Ability of the SBIV, WRAT-R,VMI and PASS for Special Education Student

Research Question 1

Will Learning disabled, mentally handicapped, gifted and parental interest groups demonstrate patterns of Binet subtest scores which have distinctive high and low subtest scores for each respective group.

Inspection of the means and standard deviations for the four groups serves to answer this question. These results are reported in Table 4.3. In Table 4.4 a summary of the rank order scores of high and low subtest scores for the special education groups is given. In table 4.4, the subtest scores are presented in descending order. For example, the LD group achieved the highest score on Pattern Analysis (51.337). This subtest is placed in rank number one in Table 4.4. The Absurdities score was slightly lower for the LD group (51.036) and it was placed in position number two. Within each group, no one subtest score deviates significantly from the mean of the group, but these scores do yield pertinent information. Table 4.4 is intended as a tool for clinical evaluation.

It should be noted that all four groups have their lowest mean subtest scores on the Memory for Sentences and Copying subtests. Bead Memory is among the low scores for MH, gifted and normal groups with scores of 34, 57 and 51, respectively. Absurdities is among the peak scores for the LD group (51), as well as for the MH (39) and normal group (54). Number Series appears in the upper rank scores for the MH (39), gifted (61) and normal (54) groups.

The gifted group's high scores are represented by the Number Series, Matrices, Quantitative and Comprehension subtests. Number Series, Vocabulary, Pattern Analysis and Absurdities are the peak scores for the normal group. The learning disabled group's high scores include Absurdities, Matrices, Memory for Objects and Pattern Analysis. The mentally handicapped group's peaks are represented by Memory for Objects, Absurdities, Number Series and Memory for Digits.

Table 4.3
Group Means and Standard Deviations for SBIV Subtests

	LD		MH		G		N	
VAR	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
SV	49.065	6.432	37.348	6.767	60.435	5.385	54.449	5.161
SC	50.407	7.740	37.295	7.213	60.439	5.441	53.725	7.087
AB	51.036	6.824	38.953	5.127	59.761	6.932	53.803	7.220
PT	51.337	6.500	38.435	8.350	59.862	7.161	54.145	6.802
CP	45.819	7.491	34.707	8.054	53.886	7.806	49.917	6.021
MT	50.776	5.559	38.206	5.918	60.781	5.085	53.073	5.419
QU	47.763	6.013	37.442	6.017	60.690	6.793	51.522	6.725
NS	50.410	5.821	38.963	4.903	61.179	4.243	54.143	4.414
BM	49.446	8.852	35.422	7.650	57.328	8.403	50.913	8.371
MS	46.813	7.194	33.848	7.501	58.250	7.989	50.174	7.081
MD	47.734	6.141	40.421	6.306	58.581	6.783	52.566	7.045
MO	50.435	6.305	40.872	6.705	55.368	8.537	52.771	5.586

TABLE 4.4
High and Low SBIV Subtest Scores for
Learning Disabled, Mentally Handicapped, Normal and Gifted Groups

	LD	MH	N	G
Rank	Subtests	Subtests	Subtests	Subtests
1	PT	MO	SV	NS
2	AB	MD	PT	MT
3	MT	NS	NS	QU
4	MO	AB	AB	SC
9	QU	SC	QU	MS
10	MD	BM	BM	BM
11	MS	CP	MS	MO
12	CP	MS	CP	CP

Low scores for the gifted group include Memory for Sentences, Bead Memory, Memory for Objects and Copying. For the normal group, low scores are represented by Quantitative, Bead Memory, Memory for Sentences, and

Copying. The learning disabled group achieve low points in Quantitative, Memory for Digits, Memory for Sentences and Copying.

In conclusion, there are patterns of high and low scores which are distinct to each educational group. However, there are very few subtests which are unique to only one group. In particular, the patterns of low scores for the four groups are quite similar. Memory for Sentences and Copying are among the low subtest scores for all four groups. Furthermore, all four groups show limited range in subtest scores. There are no subtests for any of these groups which have a score greater or less than eight points from the group mean. There are significant statistical between group differences in subtest scores, but not significant within group differences. For example, the LD group has a mean of 49.253 on all subtest scores. The subtest scores range from 45.819 to 51.337. Neither the highest score nor the lowest score deviate significantly from the mean score. The high and low scores are determined on the basis of ordinal classification only and may represent chance fluctuations. For these reasons, the confirmation of research question one is accepted.

Hypothesis 1

> Scores from the SBIV, WRAT-R, VMI and PASS will be significantly different for learning disabled, mentally handicapped, gifted, and parental interest (normal groups).

In order to determine whether scores of the administered tests are significantly different for the four groups, a oneway analysis of variance was performed on each of the variables relevant to hypothesis one. When significant F's were obtained on the anova's, appropriate scheffe post hoc comparisons were conducted. The .05 level of significance was the criterion for all tests. Oneway anova results are recorded in Appendix D. The scheffe post hoc significant differences between the four groups for the Composite and Area scores and for the specific subtests are recorded in Table 4.5 and 4.6.

TABLE 4.5
Scheffe Tests of Comparisons for SBIV Composite IQ and Area Scores

Comparisons	Subtests				
	VR	AVR	QR	STM	SBIQ
LD/MH	*	*	*	*	*
LD/G	*	*	*	*	*
LD/N	*	*	*	*	*
MH/G	*	*	*	*	*
MH/N	*	*	*	*	*
G/N	*	*	*	*	*

*significance at the .05 level is represented by *

TABLE 4.6
Scheffe Tests of Comparisons on SBIV Subtest Scores

Comparisons	Subtests											
	SV	SC	AB	PT	CP	MT	QU	NS	BM	MS	MD	MO
LD/MH	*	*	*	*	*	*	*	*	*	*	*	*
LD/G	*	*	*	*	*	*	*	*	*	*	*	*
LD/N	*	*			*		*	*		*	*	
MH/G	*	*	*	*	*	*	*	*	*	*	*	*
MH/N	*	*	*	*	*	*	*	*	*	*	*	*
G/N	*	*	*	*		*	*	*	*	*	*	

The results indicate that the Composite IQ and all four Area scores differentiate the gifted, learning disabled, normal and mentally handicapped groups.

In regards to the subtest scores which are shown in table 4.6, Number Series, Vocabulary, Comprehension, Quantitative, Memory for Sentences and Memory for Digits are significantly different for all four groups. Copying shows significant differences for all groups except for a comparison between the gifted and normal subjects. There are several subtests which discriminate between all groups except for the learning disabled and normal groups. These subtests include Absurdities, Pattern Analysis, Matrices and Bead Memory. Memory for Objects shows no significant differences between the learning disabled and

normal groups as well as no differences between the gifted and normal groups. All other comparisons are significantly different.

Oneway anovas were carried out on the WRAT-R, VMI and PASS variables to determine if these scores would show significant differences among the four groups. The results are shown in Table 4.7.

TABLE 4.7
Scheffe Tests of Comparison on
WRAT-R, VMI and PASS scores

Comparisons	Subtests										
	WR	WS	WA	PFS	GA	AT	SS	RS	PN	CON	VMI
LD/MH	*	*	*								*
LD/G	*	*	*	*	*		*	*		*	
LD/N	*	*	*								
MH/G	*	*	*		*						*
MH/N	*	*	*								
G/N	*	*	*								*

WRAT-R

As can be seen in Table 4.7, all four group means are significantly different on the Reading, Spelling and Arithmetic subtests of the WRAT-R.

PASS

On the PASS, the General Ability Scale differentiates between the learning disabled and gifted groups and the mentally handicapped and gifted groups. The Full Scale, School Satisfaction, Reading/Spelling and Confidence scale means show significant differences between the learning disabled and gifted subjects. Arithmetic and Penmanship/neatness mean scores do not discriminate among any of the groups.

VMI

For the VMI, results show significant differences between the learning disabled and mentally handicapped groups, the mentally handicapped and gifted groups and the gifted and normal groups. There are no significant differences among any other groups.

Conclusions

Based on these results, hypothesis one is accepted. As was discussed in this section, all of the SBIV area scores, the Composite IQ and several of the subtests differentiate among the four groups. Some SBIV subtests differentiate among specific groups but not among all four groups. The WRAT-R subtests differentiate among all four groups. Some PASS subtests and the VMI show limited ability to discriminate among educational groups. Based on the above confirmations of discriminative ability of the relevant tests, hypothesis one can be accepted. However, there is some refuting evidence. There are some SBIV subtests which do not differentiate among all groups. The PASS and VMI show limited ability to differentiate among educational subjects.

Hypothesis 2

> Scores from the SBIV subscales and area scores will be predictive of membership in learning disabled, mentally handicapped, gifted and normal educational groups.

To validate hypothesis 2, a linear multivariate stepwise regression was carried out on SBIV subtest scores, with the group category as the dependent variable. Table 4.8 is a summary of the results. The .05 level of significance was the criterion level for entry into the regression equation. Several discriminant functions were also done and are presented in Tables 4.9 to 4.12.

TABLE 4.8

Linear Multiple Stepwise Regression for SBIV Subtests

STEP	VAR	MULTR	R2	F	SIGF	BETA IN	CORR
1	NS	.7805	.6092	160.56	.000	.7805	.7805
2	SV	.8331	.6941	115.700	.000	.4189	.7635
3	MT	.8527	.7270	089.661	.000	.2841	.7560
4	MD	.8636	.7458	073.358	.000	.1866	.6703
5	MO					.0376	.5446
6	BM					-.0646	.5799
7	CP					.1428	.6299
8	PT					.0310	.6539
9	QU					.1271	.7401
10	AB					-.0466	.6663
11	MS					.0762	.6959
12	*SC	8762	.7677	025.330	.0218	-.0627	.6872

*Only the first four variables entered the stepwise equation at a significant level. Variables 5-12 contribute only .02 to the variance in the category variable. They do not contribute significantly to prediction of group membership.

Using a stepwise linear regression analysis, only four subtests significantly discriminated between educational group membership. These subtests include Number Series, Vocabulary, Matrices and Memory for Digits. In combination, these subtests can account for about 75% of the variance in predicting group membership. The other eight SBIV variables included in the study account for less than 1% of the variance in the educational category.

The four significant predictors from the regression equation were used in calculation of a linear discriminant function. Results are recorded in Table 4.9.

TABLE 4.9
Linear Discriminant Function of Vocabulary, Number Series Matrices and Memory for Digits for prediction of Educational Group Membership

		Predicted Group Membership							
		LD		MH		G		N	
ACTUAL GP	N	N	%	N	%	N	%	N	%
LD	93	64	63.8	4	4.3	2	2.2	23	24.7
MH	46	15	32.6	31	67.4	0	0.0	0	0.0
G	58	3	5.2	0	0.0	23	39.7	32	55.2
N	69	19	27.5	1	1.4	5	7.2	44	63.8

PERCENT OF CASES CORRECTLY CLASSIFIED 60.90

The combination of Vocabulary, Number Series, Matrices and Memory for Digits correctly classify learning disabled, mentally handicapped, gifted and normal groups for 60.90% of the cases. The mentally handicapped group has the highest correct classification rate, with about 67% of predictions being correct. Incorrect classifications primarily involve cases in the learning disabled category. Predictions of LD or normal group membership based on these four subtests are accurate 64% of the time. Distinguishing between normal and LD groups is the usual area of difficulty. The gifted population is correctly classified in only 40% of cases when these four subtests are used for prediction. The majority of cases are incorrectly classified as being in the normal group. Since these predictor variables did not produce satisfactory classification for all groups, further discriminant functions were carried out.

TABLE 4.10

Linear Discriminant Function of SBIV Subtests for Prediction of Group Membership

		Predicted Group Membership							
		LD		MH		G		N	
ACTUAL GP	N	N	%	N	%	N	%	N	%
LD	93	61	65.6	7	7.5	4	4.3	21	22.6
MH	46	2	4.3	44	95.7	0	0.0	0	0.0
G	58	8	13.8	0	0.0	26	44.8	24	41.4
N	69	25	36.2	0	0.0	3	4.3	41	59.4

PERCENT OF CASES CORRECTLY CLASSIFIED 64.66

TABLE 4.11

Linear Discriminant Function of the Stanford Binet General Abbreviated Battery for Prediction of Group Membership

		Predicted Group Membership							
		LD		MH		G		N	
ACTUAL GP	N	N	%	N	%	N	%	N	%
LD	93	55	59.1	10	10.8	4	4.3	24	25.8
MH	46	3	6.5	43	93.5	0	0.0	0	0.0
G	58	2	3.4	0	0.0	48	82.8	8	13.8
N	69	18	26.1	1	1.4	14	40.3	36	52.2

PERCENT OF CASES CORRECTLY CLASSIFIED 68.42

TABLE 4.12
Linear Discriminant Function of the Stanford-Binet Area Scores
For Prediction of Group Membership

		Predicted Group Membership							
		LD		MH		G		N	
ACTUAL GP	N	N	%	N	%	N	%	N	%
LD	93	54	58.1	6	6.5	4	4.3	29	31.2
MH	46	1	2.2	45	97.8	0	0.0	0	0.0
G	58	0	0.0	0	0.0	53	91.4	5	8.6
N	69	22	31.9	0	0.0	9	13.0	38	55.1
PERCENT OF CASES CORRECTLY CLASSIFIED 71.43									

A linear discriminant function was carried out based on all SBIV subtests. Results are found in Table 4.10. Overall, the inclusion of all subtests for prediction purposes correctly classifies groups in about 65% of cases. For the mentally handicapped group, using all twelve variables dramatically increases the percentage of correct prediction to 96 percent. The LD and normal groups achieve fairly similar prediction rates with correct classification rates of 65 and 60 percent, respectively. As in the previous discriminant function, correct classification of the gifted group occurs less than 50 percent of the time. Incorrect placement of a gifted subject into the normal category is the most common error.

Further analysis was done using the General Abbreviated Battery. This battery includes Vocabulary, Comprehension, Pattern Analysis, Quantitative, Bead Memory and Memory for Sentences. As shown in table 4.11, using these variables increases the accuracy of prediction to about 68 percent. The percentage of MH cases which are correctly classified is 94. The prediction of membership in the gifted category is accurate in 83 percent of the cases. There are 60% of LD cases which are correctly classified using the General

Abbreviated Battery as the discriminant criteria. As in previous discriminant functions, distinction between LD and normal groups is difficult to attain. Prediction of normal group membership is accurate in 52% of the cases. Subjects in the normal group are incorrectly classified as gifted 20% of the time and as LD 26% of the time.

In using the area scores as discriminants, the group members are classified correctly in 71 percent of the cases. The mentally handicapped group and gifted group are accurately classified in 98 and 91 percent of the cases, respectively. Subjects from LD and normal groups are correctly classified 58 and 55 percent of the time, respectively. The majority of incorrect classifications occur when distinctions between LD and normal groups are attempted.

In conclusion, hypothesis two can be accepted. Certain subtests and combinations of subtests are more efficient for making predictions about group membership. The area scores show the greatest efficacy for prediction. The results also show that distinctions between normal and LD groups are the most difficult to make. Predictions of membership for the mentally handicapped group are the most accurate. There is no combination of subtest or area scores on the SBIV which provides completely accurate predictions regarding group membership.

CONVERGENT VALIDITY OF THE SBIV

The convergent validity of the SBIV was examined by investigating the relationship between the Stanford-Binet IV, the Wechsler Intelligence Scale for Children-Revised and the relationship of the SBIV with the WRAT-R, VMI and PASS. In addition, pearson product moment correlations were calculated between the scores of the SBIV and the WISC-R. Results are presented in table 4.13. Pearson product moment correlations were also calculated with the SBIV and the WRAT-R, PASS and VMI. The correlation matrix for these variables is presented in table 4.14. Decimal points have been omitted. All correlations are significant at the .05 level.

TABLE 4.13

Correlation Matrix of SBIV subtests, Area scores and Composite IQ and the WISC-R subtests and IQ scores

	IN	SI	AR	WV	WC	DS	PC	PA	BD	OA	CO	MA	VIQ	PIQ	FSIQ
SV	80	78	73	85	79	67	55	58	72	62	51	41	88	74	85
SC	77	78	75	78	80	64	60	61	66	57	56	52	86	75	85
AB	72	72	67	70	72	60	63	62	67	65	44	53	78	74	80
PT	67	66	66	62	64	57	63	53	77	65	53	59	72	76	77
CP	64	62	57	65	60	58	59	63	67	66	54	54	69	73	74
MT	74	74	70	71	76	61	64	52	73	66	50	53	81	79	84
QU	74	70	70	68	70	66	53	54	67	58	43	47	78	69	78
NS	78	77	78	72	72	64	48	46	69	60	43	46	85	71	84
BM	61	58	63	58	62	54	60	50	71	61	48	44	67	71	71
MS	78	70	66	77	69	76	49	49	63	53	48	42	80	64	76
MD	68	65	67	65	64	82	47	51	65	57	50	42	73	66	72
MO	60	58	57	61	55	55	48	50	56	60	48	40	64	63	66
VR	84	83	78	85	84	70	65	66	75	67	56	54	91	81	91
AVR	76	76	74	73	75	66	68	63	82	74	60	53	83	86	88
QR	78	75	76	71	73	68	56	56	72	64	48	51	83	74	83
STM	79	74	75	78	75	77	61	60	75	67	58	51	84	79	85
SBIQ	86	84	82	83	83	76	68	66	83	74	61	59	92	87	94

TABLE 4.14
Correlation Matrix of SBIV Subtests and
WRAT-R Subtests, PASS Subtests and VMI

	WR	WS	WA	PFS	GA	AT	SS	RS	PN	CON	VMI
SV	66	66	66	* 13	29	* 01	*-07	* 12	* 08	* 10	54
SC	61	62	70	* 18	24	* 12	*-01	* 14	* 12	* 14	50
AB	57	54	63	* 10	32	* 04	*-09	* 04	* 03	* 04	60
PT	50	51	54	*-01	*14	*-09	*-04	*-11	*-02	* 12	55
CP	52	54	60	* 16	26	* 04	*-03	* 04	24	* 10	70
MT	63	59	66	* 09	27	* 01	*-01	* 06	* 07	*-03	62
QU	58	62	68	19	17	* 11	* 07	* 13	* 12	* 19	44
NS	66	68	81	* 13	27	* 14	*-14	* 10	* 01	* 15	62
BM	45	49	54	* 08	21	*-01	*-07	* 03	* 12	* 06	46
MS	61	61	60	* 15	22	* 04	* 01	* 15	* 09	* 10	40
MD	65	69	62	19	25	* 14	*-01	21	* 04	18	56
MO	53	50	50	* 12	29	* 10	*-01	* 02	* 07	* 04	45
VR	68	67	74	* 15	31	* 06	*-06	* 11	* 08	* 10	61
AVR	62	63	70	* 10	25	*-01	*-01	*-01	* 10	* 10	71
QR	62	65	73	18	23	* 13	* 02	* 12	* 10	* 19	50
STM	65	68	68	15	26	* 08	*-02	* 12	* 09	* 12	54
SBIQ	72	73	80	16	30	* 07	*-02	* 10	* 10	* 14	66

* REPRESENTS NON-SIGNIFICANCE AT THE .05 LEVEL

Hypothesis 3

The Verbal Reasoning SAS, Quantitative Reasoning SAS and Short-Term Memory SAS will correlate more highly with the Verbal IQ and Full Scale IQ of the WISC-R than with the Performance IQ.

The data relevant to this hypothesis are found in Table 4.13. The Verbal Reasoning area has a higher correlation with the WISC-R VIQ (.92) and with the FSIQ (.91) than with the Performance IQ (.81). The Quantitative Reasoning SAS has a coefficient of .83 with the Wechsler VIQ; .83 with the FSIQ and .74 with the Performance IQ. The Short-term Memory score correlates .84 with the WISC-R VIQ; .85 with the FSIQ and .79 with the Performance IQ. These results are as predicted and therefore hypothesis three is accepted.

Hypothesis 4

The Abstract/Visual Reasoning SAS will correlate more highly with scores on the Performance Scale than with scores on the Verbal Scale.

The correlation of the Abstract/Visual Reasoning Area with the Performance IQ of the WISC-R is .86. The relationship between the A/VR score and the VIQ is also strong, with a coefficient of .83. As expected, the A/VR and PIQ correlation is the higher of the two comparisons. However, the difference between the two correlations is smaller than those discussed in relation to hypothesis three. Hypothesis four is accepted, with note of the latter observation.

Hypothesis 5

The Composite score on the Fourth Edition will correlate highly with all three IQ scores on the WISC-R.

The Composite IQ shows a strong relationship with all IQ scores of the WISC-R. A highly significant correlation of .94 is found between the Composite IQ and the Full Scale IQ. The VIQ and PIQ correlate .92 and .87 with the Composite IQ respectively. Hypothesis 5 is accepted.

Hypothesis 6

The SBIV Vocabulary subtest will correlate highly with the WISC-R Vocabulary subtest.

The correlation between the two respective vocabulary tests is .85. This indicates a very significant positive relationship. On the basis of these results, hypothesis 6 is accepted.

Hypothesis 7

The Comprehension subtests from the Binet and WISC-R will be highly correlated with each other.

The correlation between the SBIV Comprehension subtest and the WISC-R Comprehension subtest is .80. Therefore hypothesis 7 is accepted.

Hypothesis 8

The Memory for Digits subtest from the Stanford-Binet will be highly correlated with the Digit Span subtest on the WISC-R.

A highly significant correlation coefficient of .82 is found when the Binet Memory for Digits and the WISC-R Digit Span are compared. Therefore, hypothesis 8 is accepted.

Hypothesis 9

The Binet subtest, Pattern Analysis, will be highly correlated with the Block Design subtest on the WISC-R

Pattern Analysis and Block Design are highly correlated, with a coefficient of .77. Hypothesis 10 is accepted.

Hypothesis 10

The Quantitative Reasoning Area, Number Series Subtest and the Quantitative subtest will correlate significantly with the WISC-R Arithmetic subtest.

The relationship between the SBIV Quantitative subtest and the WISC-R Arithmetic subtest is significant. The coefficient is .70. The correlation between Number Series and Arithmetic is higher, with a coefficient of .78. The Quantitative Reasoning SAS correlates highly with the Arithmetic subtest with a correlation coefficient of .76. Hypothesis 10 is accepted.

Hypothesis 11

The WRAT-R Reading and Spelling subtest will correlate significantly with the SBIV Composite IQ, Verbal Reasoning SAS and subtests from the Verbal Reasoning Area.

The correlation of the WRAT-R Reading subtest with the SBIV is quite high at .72. The correlation between the Reading subtest and the Verbal Reasoning SAS is slightly lower at .68. Correlations with subtests of the Verbal Reasoning area are moderate with .65 for Vocabulary; .61 for Comprehension; and .56 for Absurdities.

The correlations of the WRAT-R Spelling subtest with the Binet are very similar to those of the Reading subtest. The correlation between the Spelling subtest and the Composite IQ is .73. The VR SAS correlates .67 with the spelling subtest. Spelling and Verbal Reasoning subtest correlations are more moderate, with .65 for Vocabulary, .62 for Comprehension and .54 for Absurdities.

On the basis of these results, as reported in table 4.14, hypothesis 11 is accepted.

Hypothesis 12

The WRAT-R Arithmetic test will correlate significantly with the Binet Composite IQ, Quantitative Reasoning SAS and the Quantitative and Number Series subtests.

The correlations of WRAT-R Arithmetic with the SBIV are all highly significant. The Composite IQ correlates with the Arithmetic subtest at .80. Number Series shows the highest correlation at .81, followed by the Quantitative Reasoning SAS at .73. The correlation between the Quantitative and Arithmetic subtests is more moderate at .68. Hypothesis 12 is accepted.

Hypothesis 13

The PASS subscale scores will correlate significantly with the SBIV Composite IQ and certain subtests on the SBIV.

The correlations of all PASS subtests are all very low. Most of the correlations are not significant. Only the PASS Full Scale and the General Ability subscale show significant relationships with the Composite IQ, with correlations of .17 and .30 respectively. For the Area scores, the Full Scale

shows significant relationships with all areas except the A/VR area. Correlations range from .15 to .18. The General Ability subscale has significant correlations with all area scores, ranging from .23 to .31. The PASS Confidence subscale correlates .19 with the Quantitative Reasoning SAS.

Among the subtests of the SBIV, the General Ability subscale shows significant correlations with all subtests except Pattern Analysis and Memory for Objects. The coefficients for the Binet subtest scores and General Ability scores range from .16 to .31. The PASS Full Scale correlates significantly with Comprehension (.18), Quantitative (.19), and Memory for Digits (.19). PASS Arithmetic correlates .21 with Memory for Digits. PASS Penmanship/Neatness correlates .24 with Copying. The Confidence subtest correlates .19 with Quantitative and .18 with Memory for Digits. All other correlations between PASS subscales and SBIV subtests are not significant .

Hypothesis 13 is accepted on the basis of those correlations which are significant.

Hypothesis 14

The VMI score will correlate significantly with the Composite IQ on the Binet, the A/VR SAS and the Pattern Analysis, Copying and Matrices Subtests.

The VMI correlates .66 with the Composite IQ on the Binet. The A/VR SAS shows the highest correlation with the VMI score, with a correlation of .71. Copying has a high correlation of .70. Matrices and Pattern Analysis show more modest correlations of .61 and .55, respectively. On the basis of these results, hypothesis 14 is accepted.

Preschool Profile Analysis

Research Question 2

Will the Preschool group demonstrate a distinctive profile of high and low subtest scores? How will the results of this study compare to the results of the Technical Manual?

Means and standard deviations for the preschool group are shown in table 4.15. The scores are ranked in order from the highest to the lowest score. The Area scores and the Composite IQ are summarized in Table 4.16.

TABLE 4.15

Preschool Means and Standard Deviations for SBIV Subtests

Rank	Subtest	N	Mean	Std. Dev
1	MO	6	55.333	8.710
2	MT	3	54.000	2.646
3	MD	10	53.700	5.658
4	AB	59	53.661	5.938
5	SC	59	53.542	6.196
6	SV	60	53.050	6.748
7	QU	58	52.069	7.218
8	PT	60	51.367	7.781
9	CP	58	50.690	6.647
10	BD	59	48.254	8.425

TABLE 4.16

Preschool Means and Standard Deviations
for Area scores and the Composite

Variable	N	Mean	Std.Dev.
VR	60	107.717	12.213
AVR	60	102.433	14.074
QR	58	104.138	14.436
STM	60	107.000	15.418
SBIQ	60	107.000	13.019

The preschool group has high scores on Memory for Objects, Matrices, and Memory for Digits. However, the scores obtained are all based on a sample of ten subjects or less. Therefore, these results cannot be generalized to the preschool population. Absurdities, Comprehension and Absurdities, the three Verbal Reasoning subtests, were among the highest scores. Quantitative and Pattern Analysis were in the middle to low range of scores. The lowest

scores were obtained in Bead Memory and Copying. Copying was among the lowest scores for all four educational groups. Bead Memory was among the lowest scores for the MH, Gifted and Normal groups. The Preschool group shows strength in verbal tasks, weaker scores in Abstract/Visual skills and in the Quantitative area. Results are similar to those obtained earlier for the other educational groups.

The Preschool group achieved a mean Composite IQ of 104.3. This is slightly lower than the mean score cited in the Technical Manual of 105.3 (See Table 6.8). In this study, the highest score was in the Verbal Reasoning IQ (107.7), followed by the Short-Term Memory Area Score (107.0). The Quantitative Reasoning IQ was 104.1. The Abstract/Visual Reasoning SAS was the lowest score, at 102.3. This lower A/VR score is likely affected by the low mean copying score. In the manual, the same general ordinal trends were noted. The Verbal Reasoning Area was 109.8, approximately two points higher than in the current study. The Short-Term Memory and Quantitative scores cited in the manual were about 2 points lower than in the present study. The A/VR area was also the lowest score in the manual, with a score of 100.4. The results of this study are consistent with preschool scores as reported in the manual.

Summary

The results of this study are consistent with expectations. The discriminative validity of the SBIV was confirmed. General profiles of high and low scores were found for the learning disabled, mentally handicapped, gifted and normal groups. Copying and Memory for Sentences were low scores for all educational groups. Significant differences for the four educational groups were found for the following subtests: Number Series, Vocabulary, Comprehension, Quantitative, Memory for Sentences, and Memory for Digits. The General Abbreviated Battery and the Area IQ's were shown to be good predictors of membership in mentally handicapped groups. For learning disabled and normally achieving groups, the subtest configuration which resulted in the highest percentage of accurate predictions included Number Series, Vocabulary, Matrices, and Memory for Digits. The use of all SBIV subtests in the configuration were effective as discriminators between the learning disabled and normal groups. On the WRAT-R, all four groups showed significant differences on reading, spelling and arithmetic subtest scores. The

PASS showed limited ability to discriminate among educational groups. With the VMI, there were significant differences between the learning disabled and mentally handicapped groups, the mentally handicapped and gifted groups and the gifted and normal groups.

These results demonstrate good convergent validity of the Stanford-Binet IV. The SBIV and the WISC-R are highly related with significant correlations between all subtest scores and IQ scores. The WRAT-R subtests also demonstrate strong positive relationships with the SBIV IQ scores and appropriate SBIV subtests. The PASS, as a measure of academic self-perception, has few significant correlations with the intellectual measures on the SBIV. The VMI shows moderate to high correlations with the Composite IQ, the A/VR and Abstract/Visual Reasoning subtests.

CHAPTER FIVE
DISCUSSIONS and CONCLUSIONS

Introduction

The Stanford-Binet Fourth Edition was published in 1986. As a new test, proof of its technical merit is required. The purpose of this study was to consider the validity of the new Stanford Binet. "In order to demonstrate construct validity, we must show not only that a test correlates highly with other variables with which it should theoretically correlate, but also that it does not correlate significantly with variables from which it should differ" (Campbell and Fiske, in Anastasi, 1982). This study focused on the convergent validity of the SBIV. A comparison was made with one of the most widely used intelligence tests, the Wechsler Intelligence Scale for Children-Revised. There are several subtests on the Stanford-Binet IV which appear to have similar counterparts on the WISC-R. The actual relationship between these specific subtests was of particular interest. The relationship of the new Binet, the WRAT-R, PASS and VMI was also of interest. These tests represent measures of academic achievement, scholastic self-concept and visual-motor skills respectively.

The ability of the Binet IV to discriminate among special educational groups such as the mentally handicapped, LD , normal and gifted groups would confirm the discriminative validity of the test. Therefore, information regarding the discriminative validity of the SBIV has important implications for educational psychologists. In general the results of the study support the discriminative and convergent validity of the Stanford-Binet Fourth Edition. Specific results are discussed within each relevant section.

Profile Analysis of Special Populations

Learning Disabled

The rank order of the LD group from highest to lowest is as follows: Pattern Analysis, Absurdities, Matrices, Memory for Objects, Number Series, Comprehension, Bead Memory, Vocabulary, Quantitative, Memory for Digits, Memory for Sentences and Copying. The results show that the LD children

achieve high scores on Pattern Analysis and Matrices which are both components of the Abstract/Visual Reasoning area. Absurdities is in the second highest position. This test involves verbal explanation of visual incongruities and requires both verbal and visual processes. As a visual task, it requires skills which are common to the Pattern Analysis and the Matrices task. Memory for Objects involves short-term memory for familiar concrete objects and the correct sequential presentation of these objects. All top four scores for the LD group involve visual acuity and discrimination.

The middle scores are represented by Number Series, Comprehension, Bead Memory and Vocabulary. Comprehension and Vocabulary are included in the Verbal Reasoning area. The ability to define words, express oneself and use common sense to solve social dilemmas are measured by these tasks. Verbal skills appear to be less well developed than visual skills for the LD group. The Number Series subtest was at position number five in ranking, and requires the evaluation of the numerical pattern and the application of mathematical rules. A reduced score on this test is consistent with the academic weakness in arithmetic for some LD children. Bead Memory has been discussed as a measure of simultaneous processing and visual short-term memory (Osberg, 1986; Thorndike, Hagen and Sattler, 1986b). These tests are relatively or moderately difficult for LD children.

The lowest scores for the LD group are on Quantitative, Memory for Digits, Memory for Sentences and Copying. The Quantitative subtest involves the solving of simple arithmetical problems and understanding of numerical operations. Memory for Digits requires repetition of a given sequence of numbers using short-term memory and attentional and concentration skills. Memory for Sentences is also among the low scores. A weakness in this area may reflect attentional and concentration difficulties as in the Memory for Digits Subtest. This task requires memory for verbal semantic material. The verbal nature of the task may make it difficult for the learning disabled group. Memory for Sentences and Copying are among the lowest scores for the learning disabled group but also for the normal, gifted, and mentally handicapped groups.

The general pattern of scores found for the LD group shows strength in visual tasks, less developed abilities for verbal tasks and low scores in

mathematical and short-term auditory memory tasks. Weaknesses in the latter task may reflect poorly developed attentional and concentration skills.

The results presented here are consistent with the recent literature on WISC-R profile analysis of LD children. Sattler (1982) reviewed the scores of reading disabled children who were given the WISC-R. A summary of the results from thirty studies showed Performance subtests such as Picture Completion, Pattern Analysis , Block Design and Object Assembly to be the highest mean scores. These WISC-R tasks are based on visual processes as are the high scores on the SBIV. The WISC-R Comprehension and Vocabulary subtests were among the middle scores for the reading disabled children. The SBIV Comprehension and Vocabulary subtests are ranked as middle scores in the current study. Sattler found the four most difficult WISC-R subtests to be Arithmetic, Coding, Information, and Digit Span which form the acronym ACID. The Quantitative subtest and the Memory for Digits subtest are low scores for the LD group in the current study. They can be accepted as the most equivalent SBIV measures of the "A" and "D" components of the ACID tests. These results have been replicated in other studies (Fowler, 1986; Galvin, 1981; Kaufman, 1979; Kaufman and McLean, 1986; McManis, Gigley, Richert, and Fabre, 1978; Moore and Wielan, 1981; Mueller et al.,1986; Naglieri, 1980; Naglieri, 1985; Schiff, Kaufman and Kaufman, 1981; Sexton and Street, 1985; Smith, 1978; Stevenson, 1980; Strichert and Love, 1979; Ysseldyke, Algozzine, Shinn and McGue, 1982).

Factor analytic studies of LD groups have shown the Perceptual Organization (PO) tasks to be relative strengths. Verbal Comprehension tasks are more difficult for them. Subtests which load on the Freedom from Distractibility factor receive the lowest scores. (Galvin, 1981; Kaufman, 1981; Kaufman and McLean, 1986; Kavale and Forness, 1984; Mueller, Dennis and Short, 1986). This general pattern was replicated in this study.

Bannatyne's recategorization system has also been used to investigate the typical WISC-R profile of the LD population. Individuals with learning disabilities often have a WISC-R pattern of Spatial > Conceptual > Sequential (Clariozo and Bernard, 1981; Dudley-Marling, Kaufman and Tarver, 1981; Kaufman, 1979; Kaufman, 1981; Sattler, 1982; Smith, Coleman, Dolecki and Davis, 1977; Vance and Singer, 1979). These general trends were also found in this study.

These data are generally consistent with the results of a recent study by Meloff (1986). In her study, the learning disabled group's high scores were represented by Absurdities, Matrices, Number Series, Memory for Digits and Memory for Objects. Absurdities, Matrices, and Memory for Objects were found to be high scores in the present study as well. The LD groups in both studies have strengths in tasks which are based on visual skills. There is one discrepancy between the two studies which is very surprising. In Meloff's study, Memory for Digits appeared as a high score; in the present study it appeared as a low score. As was mentioned earlier, the literature suggests that Memory for Digits is likely to be a low score for the LD group. The discrepancy may be due to a difference in sample size, with Meloff's study having a much smaller sample. Recent studies suggest that the learning disabled population does not exist as a homogeneous entity but is composed of several heterogeneous subgroups with varying patterns of abilities (McKinney, 1984; Snow, Cohen and Holliman, 1985). An attention deficit may be present in many of the individuals in the current study but not in Meloff's group.

In Meloff's study, the low mean scores were found for the learning disabled group in Copying, Quantitative, Memory for Sentences, Bead Memory and Vocabulary. These results are similar to those found in the current study.

There are two important considerations in using profile analysis for the diagnosis of learning disabilities. First of all, Sattler (1982) states that "there is no one particular WISC-R pattern that is reliably diagnostic of learning disability. The WISC-R should be used to assess the child's intelligence and patterns of cognitive efficiency; it should not be used in the absence of other tests and information to make a differential diagnosis" (Sattler, p. 402). These considerations are also applicable to the Stanford-Binet Intelligence scale. Each interpretation of test scores must be done in an ipsative manner using integration of background information, clinical observations, and supplementary test data. Secondly, current research suggests that the LD population does not represent a distinct homogeneous group, but is rather, heterogeneous in nature (McKinney, 1984). Snow, Cohen and Holliman (1985) found six distinct subgroups when the scores form the WISC-R were factor analyzed. These subgroups include individuals who are low achieving in all academic areas, have language deficits, attentional deficits, show discrepancies in visual and verbal tasks or have patterns of low average to average abilities.

Mentally Handicapped

The rank order of the Binet IV subtests from highest to lowest is as follows: Memory for Objects, Memory for Objects, Number Series, Absurdities, Pattern Analysis, Matrices, Quantitative, Vocabulary, Comprehension, Bead Memory, Copying and Memory for Sentences.

The pattern of high and low scores is generally as expected, with strengths in perceptual tasks and reduced scores in verbal and academic tasks. Tasks which are concrete in nature also appear to be stronger skills for this group. Abstract skills are less well developed. Overviews of the WISC-R scoring patterns for the MH population show a similar profile to that found in the present study. The best achievement occurs on tasks which required perceptual skills; weaknesses occur on tasks which require verbal comprehension and expression (Kaufman, 1979; Kaufman and Van Hagan, 1977; Mueller, Dash, Matheson and Short, 1984; Mueller, Dennis and Short, 1986; Naglieri, 1980; Naglieri, 1985; Sattler, 1982). The typical profile of subtest scores is quite similar for both the MH and LD groups. In this study both these groups achieved higher scores on subtests such as Pattern Analysis and Matrices (Perceptual tasks) and reduced scores on subtests such as Vocabulary, Quantitative and Comprehension (school related tasks).

The results in this study correspond to those of Meloff (1987). In Meloff's study, the mentally handicapped group obtained high scores in Memory for Objects, Absurdities, Pattern Analysis and Memory for Digits. The current study also includes Number Series and Matrices as high scores. Low scores in the research of Meloff included Memory for Sentences, Copying, Vocabulary and Quantitative subtests. The current study shows the Vocabulary and Quantitative subtests to be in the middle range of scores: Bead Memory and Comprehension are shown as low scores.

Gifted

For the gifted group, the scores on subtests appear in the following rank order (from highest to lowest): Number Series, Matrices, Quantitative, Comprehension, Vocabulary, Pattern Analysis, Absurdities, Memory for Digits, Memory for Sentences, Bead Memory, Memory for Objects, Copying.

The results are as expected, with the high scores on subtests which have the highest loadings on the general ability factor (Number Series, Matrices, Quantitative, and Vocabulary). Abstract reasoning and nonverbal reasoning appeared as strengths for the gifted group with high scores on Matrices and Number Series. Two of the Verbal Reasoning subtests, Comprehension and Vocabulary, are also high scores for the gifted group. Well developed verbal facility has been reported for the gifted. Factor analytic studies have shown that the Verbal Comprehension Factor of the WISC-R usually has the highest mean score although the Perceptual Organization Factor may sometimes be the strongest skill (Kaufman, 1981; Mueller, Dash, Matheson and Short, 1984; Mueller, Dennis and Short, 1986). The gifted group did less well on memory tasks. This finding is also consistent with recent research. In an analysis of WISC-R scores, Mueller, Dennis and Short (1986) found the gifted to obtain the lowest scores on short-term memory tasks which require sequential processing and attentional and concentration tasks.

In Meloff's study, the gifted group show peak scores in Quantitative, Number Series, and Vocabulary subtests. This is essentially the same as in the present study. This study adds Matrices and Comprehension as high scores. Low scores in both studies are represented by Copying, Memory for Sentences, and Memory for Objects. Additions to the low scores from the current research include Memory for Digits and Bead Memory.

Normally Achieving

For the normally achieving group the rank order of scores from highest to lowest is as follows: Vocabulary, Pattern Analysis, Number Series, Absurdities, Comprehension, Matrices, Memory for Objects, Memory for Digits, Quantitative, Bead memory, Memory for Sentences, and Copying.

The normally achieving group show peak scores in a wide range of tasks. These subtests represented 3 out of 4 Reasoning Areas and represent both crystallized and fluid abilities. This suggests that normally achieving children have a broad base of skills. Mueller et al. (1984) found that normal children achieve relative strengths on perceptual tasks. In this study, high scores were achieved on Pattern Analysis and Absurdities, both of which

require visual processing. Vocabulary is a high score for the normal group. This serves as a useful distinction between the LD and the normal group.

In comparing this study with the work of Meloff, the following subtests appear as high scores for both studies: Number Series, Vocabulary and Pattern Analysis. Meloff also listed Matrices as a high score. Matrices is ranked in the middle range of subtest scores in this study. Memory for Digits is listed as a high score in Meloff's study, but as a low score in the present study. This inconsistency may be due to sample differences. Perhaps the individuals in Meloff's study had more highly developed memory strategies. The differing results may also be an artifact of sample size differences. Low scores in both studies are represented by Memory for Sentences, Bead Memory and Copying. Quantitative is listed in the current study as an additional low score.

Summary

There do appear to be specific patterns of high and low scores for each educational group. The LD group shows strength in perceptual tasks, less developed abilities in verbal skills and weaknesses in attentional tasks. The MH group also shows strength on visual tasks thus showing some similarity with the LD group. The gifted group have well developed conceptual abilities and verbal skills and are proficient in academic tasks. The normal group is somewhat like the MH and LD groups with some high scores on perceptual tasks. These results are consistent with the current publications on WISC-R profile analysis. They also correspond fairly closely with recent work on the SBIV completed by Meloff. It should be noted that Memory for Sentences and Copying are low scores for all of the four groups. It is suggested that scoring procedures are overly stringent for the Copying subtest.

Pattern analysis may be used to suggest hypotheses as to classification or confirm some tentative ideas. This method should not be used in isolation to make classification decisions. Use of supplementary tests, and integration with other pertinent clinical information is necessary.

Analysis of Variance Findings

SBIV

In the current study, the discriminative validity of the SBIV was examined by comparing the mean scores of all SBIV subtests and the Area scores to determine if there were significant differences among the four educational groups. Most of the subtests and area scores do discriminate among all educational groups.

All area scores, the Verbal Reasoning area, the Abstract/Visual Reasoning area, the Quantitative Reasoning Area and the Short-Term Memory Reasoning area served to discriminate among all four groups. These results are slightly different from a previous study. Meloff reported significant differences for all four groups on the Binet IV Composite IQ and Abstract/Visual Reasoning IQ. Significant differences were recorded for Verbal Reasoning, Quantitative Reasoning and Short-Term Memory for most groups, but there was no significant difference between learning disabled and normal subjects with these scores. In contrast, the present study shows significant differences between all groups. The discrepancy may be due to differences in sample size.

The technical manual also showed that the SBIV can be used as a discriminative tool. The manual compares the scores of gifted, learning disabled and mentally handicapped students. The means of the gifted examinees on all of the areas and the Composite of the SBIV were significantly above the means of the standardization sample, and the means of the learning disabled and mentally handicapped were significantly below the means of the standardization sample. The means on all the areas and Composite of the SBIV for the mentally handicapped examinees were significantly lower than those for the learning disabled examinees and the SBIV can reliably discriminate between these two groups of exceptional students.

In this study, the subtests which showed significant differences between all four educational groups included Vocabulary, Comprehension, Quantitative, Number Series, Memory for Sentences and Memory for Digits. In Meloff's study, only the Vocabulary subtest reported significant differences for all four groups. This discrepancy is probably best explained by the increased sample size of the present study over Meloff's study. Number Series, Vocabulary,

Comprehension and Quantitative subtests are highly correlated with *g* , with loadings of .79, .76, .71 and .78 respectively. The mean scores of subtests that are highly associated with general or global intellectual ability are significantly different for all four educational groups. This finding is consistent with the predictions of Mueller et al. (1984) who stated that IQ score or general ability is the best discriminator of educational groups.

Memory for Sentences and Memory for Digits also discriminate among all four educational groups. Short-term memory differences may be interpreted in several ways. Decreased scores (as for the MH group) or increased scores (as for the gifted group) may mean that attentional factors are important. Smith (1983) suggested that LD children are less able to sustain attention as a task is being presented and do not attend to the critical distinctive features of a task. Mentally handicapped may show weakness in the ability to attend to the appropriate details of a presented stimuli (Westling, 1986).

The varying scores in Memory for Sentences and Memory for Digits may be related to differences in short-term memory capacity. Miller (1956) proposed that the capacity of short-term memory ranged from five to nine items, with the average capacity being seven items. An item is a chunk of information ranging from a single letter to an idea expressed by a paragraph. It is possible that educational groups may vary in their short-term memory capacity. Furthermore groups may vary in their ability to organize several items into a single unit or chunk. For example, the mentally handicapped group seem to be less able to organize and synthesize information (Westling, 1986).

Varying usage of cognitive strategies or mnuemonics may also be related to the distinctive memory scores between the groups. Rehearsal of information as it is presented, use of association strategies, chunking of information, or use of visual imagery all serve to increase the amount of material remembered. It is likely that high achievers utilize these strategies often, normal children use them occasionally and LD and MH children are not adept at using them. For example, Memory for Sentences is a highly verbal task. Organizing the information based on knowledge of typical syntax and grammatical structure would tend to increase the obtained score. This should give the normal and gifted groups an advantage. A review of studies by Smith (1983) indicates that LD children are less likely to verbally rehearse to remember items. Ellis and

Hunt (1983) suggest that MH children do not rehearse presented material or use paired associations with familiar words.

In this study, Copying shows significant differences for all group comparisons except between the gifted and normal groups. In Meloff's study, Copying showed significant differences between all four groups except between the learning disabled and normal groups and between the gifted and normal groups. This slight difference is results is most likely due to the differences in sample sizes. A delay in visual-motor integration skills has been described as characteristic of the LD group (Koppitz, 1981). Janzen et. al (1983) describe visual-motor tasks as adequate discriminators of LD and normal educational groups. However, some studies suggest, as does Meloff's, that visual motor integration tasks are not significantly different for LD and non LD samples (McManis et. al, 1978; Ysseldyke, Algozzine, Shinn and McGue, 1982). These contradictory difference can also be explained based on the heterogeneous nature of the LD population. It is likely that there is a distinctive subgroup of LD individuals which show weakness on visual-motor tasks (Leton, Meyamoto and Ryckman, 1987). Further investigation into the discriminative ability of visual motor tasks is warranted.

Several subtests distinguish between all groups except for the learning disabled and normal groups. These include Absurdities, Pattern Analysis, Matrices and Bead memory. Bead Memory was also unable to discriminate in Meloff's study between these two groups. Pattern Analysis, Matrices and Bead memory were reported in Meloff's study as unable to discriminate between the LD and normal groups as well as the gifted and normal groups. An increase in sample size in the present study most likely accounts for the improved discrimination ability of these subtests. It is often difficult to distinguish between the LD and normal population. Part of the problem is lack of a clear definition for learning disabled. Epps, McGue, and Ysseldyke (1982) asked eighteen judges to discriminate between an LD and a non LD, low achievement group. The results showed the judges to be extremely inaccurate and in little agreement with each other. They concluded that it is unlikely that school personnel can accurately identify LD students when they are compared with low achievement students. Ysseldyke, Algozzine, Shinn and McGue (1982) compared the scores of LD and low achievement students on a psychoeducational battery and found that an average of 96% of the scores were

within a common range and the scores on many subtests were identical. Their conclusion is that there is considerable misclassification in identification of LD students. The definitions and the criteria used to identify LD students are numerous. There is considerable controversy as to which, if any, of these definitions is correct. Epps, Ysseldyke and Algozzine (1985) found three general conceptualizations of learning disabilities. The first way of distinguishing LD students is to note any discrepancies between achievement and ability. Secondly, LD groups have been characterized as showing achievement deficits and showing overall low scores in academic areas. The use of pattern analysis, in which the pattern of high and low scores are used as indicators, has also been advocated. Using factor analytic techniques, Epps et al., (1985) found that low achievement contributed most highly to correct classification procedures and was four times more effective than the ability-achievement paradigm. Scatter did not contribute independently to classification. It may be, with such confusion over definition and classification systems varying in different regions that LD students are being misclassified.

In summary, research shows that the criteria for LD groups vary and that distinctions between LD groups and low achieving children are difficult to make. These two factors may explain some of the lack of significant differences on the SBIV subtests.

Summary

The SBIV is a test of good discriminative ability. Tasks which load highly on the general ability factor distinguish among the groups particularly well. Subtests which fall under the Crystallized Ability factor (academic tasks) also differentiate between the learning disabled, mentally handicapped, normal and gifted groups. Auditory memory tasks also differentiate well, perhaps based on inter- group differences in attentional skills, STM capacity and efficient use of mneumonics. The most difficult groups to distinguish are the LD and normal groups. Normal and gifted groups are also difficult to distinguish on some subtests. The parental interest group may represent high achievers.

Discriminative Ability of the WRAT-R

The WRAT-R is often used to provide an estimate of academic ability of a student. The results of this test show that all four groups are significantly different on mean scores for the Reading, Spelling and Arithmetic subtests. In Meloff's study all four groups were significantly different on the WRAT-R Arithmetic and Reading subtests. WRAT-R spelling differentiated between all the groups except between learning disabled and normal subjects. Again, differences between the two studies can be attributed to differences in sample size.

The results are consistent with findings in the literature. The Wide Range Achievement Test (WRAT) has been shown to discriminate between learning disabled and normally achieving students (Janzen et. al, 1983). Prasse, Siewert and Breen (1983) found the WRAT scores of an LD sample to be significantly lower than the scores of normally achieving students.

Discriminative Ability of the PASS

Analysis of the Pass scores shows that the General Ability scale differentiates between the learning disabled and gifted groups and the mentally handicapped and gifted groups. Distinctions between these groups are usually notable on achievement measures and it is not surprising that perceptions of general ability are also significantly different. The Full Scale, School Satisfaction, Reading/Spelling and Confidence scales show significant differences between the learning disabled and gifted subjects. A reduced general sense of competency and self esteem for LD subjects has been reported by several researchers (Chapman and Boersma, 1979; Kolligian and Sternberg, 1987; Schiff et. al, 1981; Ysseldyke et al, 1982). The results of this study are consistent with the previous findings. Arithmetic and Penmanship/Neatness do not discriminate among any of the groups. It is possible that these are areas in which self-concept deficits are not as predominant. Referrals to the clinic may more often involve reading difficulties or overall ability.

Meloff found slightly different results. There were significant differences reported for mentally handicapped and normal subjects on General Ability. Mentally handicapped groups and normal groups showed significant differences on the Reading/Spelling subscale. There were no significant differences reported for the four groups for the Full Scale, Arithmetic, Penmanship/Neatness of Confidence scales. Slight differences may be accounted for by considering sample size differences.

The PASS is less able to discriminate between groups than the SBIV. The manual reports that subscales of the PASS are not highly correlated with intellectual measures.

Discriminative Ability of the VMI

For the VMI, there were significant differences between the learning disabled and mentally handicapped groups and the gifted and normal groups. There were no significant differences between the LD and normal groups. Meloff (1987) found no significant differences between learning disabled and mentally retarded or between the gifted and normal populations based on VMI scores. There were significant differences between the learning disabled and normal groups in her study.

Group Classification by Linear Regression and Discriminant Function Procedure

Prediction and discrimination of group membership with the SBIV was studied by conducting a stepwise regression analysis and a series of linear discriminant functions. The findings indicate that subtests which have the highest loadings on the general ability factor are the most predictive of membership in any of the four educational groups. For example, Number Series, Matrices and Vocabulary were all significant predictors of educational category.

Memory for Digits also entered the stepwise equation at a significant level. As discussed earlier in this chapter, differences in mean scores on memory tasks may be related to inter-group differences in attentional skills, STM capacity or use of cognitive strategies.

The results provide some support for the utility of the hierarchical theoretical model of the SBIV. The importance of *g* is substantiated given that three out of four predictor variables load highly on the General Ability Factor. Furthermore, it is notable that subtests from both the Crystallized (Number Series, Vocabulary) and Fluid-Analytic (Matrices) Factors serve as effective predictor variables. The inclusion of Memory for Digits demonstrates that a Short-Term Memory factor is also a valid and necessary second level factor.

In summary, these results suggest that the best prediction of educational categories require subtests that are strongly associated with general ability, represent skills which require Crystallized, Fluid Analytic and Short-Term Memory Skills and are content specific in nature. Results provide preliminary substantiation of the theoretical model of the new Binet.

When the four significant predictors from the regression equation were used in calculation of a linear discriminant function, the prediction rates were less than adequate. Use of the combined scores on the Vocabulary, Number Series, Matrices and Memory for Digits subtests correctly classified only 61 percent of the subjects. There was a particular weakness in distinguishing between the gifted and normal population. The gifted subjects were correctly classified only 40% of the time. Therefore, further linear discriminant functions were carried out.

In a further analysis, all SBIV were used for a linear discriminant function. This increased the overall prediction rate to 65%, but the distinction between some educational groups was still inadequate. Of the 46 MH subjects, 44 or 96 percent were correctly classified. Using all twelve subtests did not adequately distinguish between the gifted and normal subjects. Of the 58 gifted individuals in this study, 26 (44.8%) were correctly identified as gifted, 24 (41.4%) were inaccurately categorized as normal and 8 (13.8%) were described as learning disabled. For the normal group, using these subtests distinguished them relatively well from the LD group. Using this combination of subtests, 64% of normal subjects were correctly identified. This was the highest classification rate obtained in all linear discriminant functions. It is obvious that all subtests do not discriminate equally well among all groups.

The General Abbreviated Battery was used to discriminate among the educational groups. This battery includes Vocabulary, Comprehension, Pattern Analysis, Quantitative, Bead memory and Memory for Sentences. Based on this

combination, the overall accuracy of prediction increased to about 68 percent. The correct prediction of membership in the MH group was quite high, at about 94 percent. This battery showed great efficacy for correct classification of the gifted group, with 83 percent accuracy. Subjects of the LD group were correctly classified in about 60% of the cases; Subjects from the normal group were accurately placed about 52% of the time. Incorrect classification usually involved confusion of LD and normal groups.

The General Abbreviated Battery seems the best current choice for making classification decisions, particularly for the MH and gifted groups. It is possible that there is a combination of subtests which would be even more efficient. For the LD and Normal group Number Series, Matrices, and Memory for Digits may be useful supplements to this battery. Evidence from the regression analysis done in this study suggests that they serve as good discriminators for these groups.

The area scores were entered separately into a linear discriminant function. These scores are measures of broader cognitive skills and therefore it is not surprising that they are more potent general discriminators. As with the subtests, the MH group is the easiest to classify. Forty-five out of 46 subjects were correctly identified as mentally handicapped. Among the gifted subjects, 91 percent were accurately identified as gifted. When trying to correctly identify either LD students or normal students, the accuracy rates fall at about 58 percent and 55 percent respectively. These level of accuracy are adequate, but not quite as adequate as when all subtests are used or the subtests from the regression equation are used. The best combination of subtests used to discriminate between educational groups may differ depending on which groups are compared.

Convergent Validity

Stanford-Binet vs the WISC-R

Results of this study show that SBIV and the WISC-R are highly related. They appear to be measuring similar constructs. The correlation coefficient between the Stanford-Binet Composite IQ and the WISC-R Full Scale IQ was .94. The Composite IQ correlated .92 with the Verbal IQ and .87 with the Performance IQ on the WISC-R. The correlations in this study are higher than those reported by Thorndike et al. The coefficient given in the manual between the SBIV Composite and the FSIQ was .83. The Composite correlated .78 with the WISC-R Verbal scale and .73 with the WISC-R Performance scale. The correlations of the Area scores ranged from .64 to .68 with the VIQ; and from .63 to .67 with the PIQ. This study supports predictions which were made in the manual. The results of this study show some similarity with recent work by Rothlisberg (1987). Carvajal and Weyand (1986) found the Composite IQ of the General Abbreviated Battery to correlate .78 with the WISC-R Full Scale IQ. This correlation is lower than that found in the current study. Differences may be due to a small sample size. Their sample consisted of 23 subjects. Nevertheless, good convergent validity is demonstrated both in this study and in the present one.

All subtest intercorrelations on the SBIV and WISC-R were significant. Furthermore, subtests which appeared to be measuring similar tasks correlated very highly. For example, the SBIV Vocabulary subtest correlated highly with the WISC-R Vocabulary subtest. The correlation between the two respective Comprehension subtests was significant.

Convergent Validity of the WRAT-R

In summary, the results of this study are consistent with previous research and demonstrate a significant relationship between these two academic and intellectual measures. These results also affirm the convergent validity of the SBIV.

Convergent Validity of the PASS

These low correlations are consistent with the findings in the PASS manual that self-concept scores are weakly associated with IQ scores (Boersma and Chapman, 1985). Meloff (1987) found that the PASS General Ability scale correlated with the A/VR area, Pattern Analysis and Copying subtests (.31 to .39). No other significant correlations were found.

VMI

In this study, The VMI showed very high correlations with Copying (.70), the Abstract/Visual Reasoning Area (.71) and the Composite IQ (.66). Correlations between the VMI and the other SBIV subtests and Area scores were more modest, ranging from .40 (Memory for Sentences) to .62 (Matrices and Number Series). All correlations were significant.

These results are much higher than those found in previous studies. Beery (1967) reported correlations between the VMI and mental age on the Primary Mental Abilities test of .59 for grade one children; .37 for grade four children and .38 for grade seven children. Wright and DeMers (1982) found that the VMI correlated .58 with Verbal IQ, .66 with Performance IQ and .66 with the Full Scale IQ.

Meloff (1987) found moderate correlations between the Binet and the VMI. There was a low, but significant correlation with the Composite score (.28). The VMI correlated with Copying at .61 and .41 with Matrices. There were no significant correlations with the Abstract/Visual Reasoning score (.24) or with the Pattern Analysis subtest score (.17).

Preschool Group

The preschool group obtained their highest score on the Vocabulary test, indicating that they have well developed language skills. The Pattern Analysis and Quantitative subtest were in the middle to low range of scores. Bead Memory and Copying were the lowest mean scores for the preschool group.

The scores obtained for the preschool group were consistent with studies

reported in the Technical Manual. In both the current study and in the manual, the highest score was obtained on the Verbal Reasoning Area, followed byShort-Term Memory and Quantitative Reasoning . The Abstract/Visual Reasoning area represented the lowest mean score.

The sample for the preschool group was quite small. This section is regarded as a preliminary study on preschool children and further research is required.

Future Research

The factor analytic studies reported in the manual do not support the subtest divisions (Slate, 1986; Vernon, 1987). There is a need for further factor analysis to clearly delineate what each subtest is measuring. The results of a survey of student clinicians showed that 59% found it difficult to know what the subtests were measuring on the SBIV (Janzen, Boersma, and Krausher, 1987). Furthermore, since different combinations of subtests are given to different age groups, investigation of the factor analysis across a broad age range is necessary. Work in this area is currently being carried out (Janzen, personal communication, 1987).

In this study, Equation Building, Paper Folding and Cutting and Verbal Relations were eliminated from the analysis due to limited sample sizes. Further research is needed which includes these subtests and considers their discriminative ability. It is also necessary to correlate these subtests with other related measures such as subtests and IQ scores on the WISC-R.

The content of the SBIV is limited for preschool children. Only eight subtests are generally given to children age two to six. Furthermore, the basal levels are not adequate for low functioning preschool children below four years of age. These problems have been noted in the literature (Telzrow, 1987; Rothlisberg, 1987). Further work is needed to validate the utility of the SBIV with preschool populations. A sample which includes a wide range of scores is necessary. Comparisons of the SBIV with similar tests such as the WPPSI or McCarthy would be beneficial.

Additional information on adult profiles on the SBIV is also required. As Vernon notes, there are no age norms for adults above the age of 23. Comparisons with the WAIS-R would serve to validate the utility of the Binet.

There is some ongoing research in this area which will add to the current information on this age range (Janzen, personal communication, 1987)

Further work is needed to investigate the affect of American content on the scores of Canadian children and to develop a Canadian norming sample. There is some cultural bias in the SBIV. Several items are American in content. For example, no allowance has been made in the Quantitative subtest for conversion to metric units of measurement. Item number eight in the Vocabulary subtest requires the identification of the American flag. The Quantitative subtest also requires the identification of American coins. The problem of American bias in the Binet has also been noted by Vernon (1987) and Meloff (1987).

Summary

The section on discriminative validity indicates that the Binet test is able to discriminate among LD, normal, gifted and mentally handicapped students. Meloff (1987) also found that the Binet could discriminate among these groups. The results of this study also corroborate statements in the manual that the SBIV can discriminate between learning disabled and mentally retarded populations and between gifted and normal populations.

Analysis of subtest profiles of educational groups indicates that each group has a general pattern of high and low scores. The learning disabled and mentally handicapped groups both show relatively high scores on visual and perceptual tasks with concomitant reduced scores on verbal tasks. The gifted group show strengths on verbal and academic tasks. The normal group showed high scores on perceptual tasks. These general patterns follow the trends demonstrated in the literature for WISC-R profiles for special educational groups. Use of profile analysis may be used as an aid in classification procedures and in generating or substantiating hypotheses. Each case must be treated as an individual case. Classification decisions should not be made solely on the basis of intellectual test scores.

The mean subtest scores are significantly different for the four groups of this study. Number Series, Vocabulary, Comprehension, Quantitative, Memory for Sentences and Memory for Digits display significant differences for all four groups. These may be useful in a generalized battery for any type of

classification decision. It is notable that some of these subtests load highly on the general ability factor. Discriminating between the learning disabled and normal groups appears to be quite difficult. In fact, the test scores from Absurdities, Pattern Analysis, Matrices and Bead Memory show significant differences between all groups except the learning disabled and normal groups. In choosing which subtests to give to a client who may be learning disabled, it is recommended that some from the first list be given priority. The WRAT-R is a useful supplement to a battery. All three subtests of the WRAT-R discriminate among all four educational groups.

The subtests and area scores of the SBIV can be used to predict educational classifications. The combination of subtests or area scores which are most beneficial vary from group to group. For correct classification of an individual who is learning disabled, utilizing all 12 subtests of this study provided the best prediction rate. For a normally achieving child, the best combination of subtests found in this study included Vocabulary, Number Series, Matrices and Memory for Digits. The area scores and the General Abbreviated Battery served as the best predictors of the MH and gifted group. It is possible that there is a combination of subtests which may serve as a slightly better predictor. The results in this study indicate that the Matrices, Number Series and Memory for Digits subtests and the subtests of the General Abbreviated Battery would serve as a useful combination. Given the results of the anova, Bead Memory should be reevaluated as part of a general battery. Absurdities, Pattern Analysis and Memory for Objects do not appear to discriminate well between normal and LD groups. Memory for Objects and Copying also fail to discriminate between normal and gifted groups. Each testing battery battery must be guided by the nature of the referral question. At this point, the SBIV appears to provide accurate predictions about 70 percent of the time. Additional information is necessary to make classification decisions.

Convergent validity of the Stanford-Binet Fourth Edition was substantiated in this study. The SBIV scores correlated significantly with the subtest and IQ scores on the WISC-R. The Composite correlated .94 with the FSIQ; .92 with the VIQ and .89 with the PIQ. Results show that the Binet is a highly verbal test. Correlations between the WRAT-R and Binet were all significant. The correlations of the VMI with the SBIV were more moderate. The correlations of the PASS subtests were all quite low.

Scores from the preschool groups were consistent with those reported in the manual. There was some evidence of a distinctive profile of high and low scores, although results are very preliminary. Further research is needed.

Overall the results show the new Binet has good convergent validity. This test is able to discriminate among learning disabled, mentally handicapped, gifted and normally achieving individuals.

REFERENCES

References

Anastasi, A. (1982). Psychological Testing. Fifth Edition. New York: Psychological Corporation.

Bannatyne, A. (1974). Diagnosis: A note on recategorization of the WISC-R scaled scores. Journal of Learning Disabilities, 7, 272-273.

Barnard, J. W. (1987). Considerations Regarding the Use of the New Stanford-Binet. The Alberta School Psychologist, 7(1), 16.

Beery, K. E. (1967). Visual-Motor Integration. Monograph. Chicago: Follett Publishing Company.

Beery, K. E. (1982). Revised Administration, Scoring and Teaching Manual for the Developmental Test of Visual-Motor Integration. Toronto: Modern Curriculum Press.

Boersma, F. & Chapman, J. (1985). Manual for the Student's Perception of Ability Scale. (rev. ed.) Edmonton: University of Alberta.

Bossard, M. D. and Galusha, R. (1979). The Utility of the Stanford-Binet in Predicting WRAT Performance. Psychology in the Schools, 19, 297-304.

Brock, H. (1982). Factor structure of intellectual and Achievement Measures for Learning Disabled Children. Psychology in the Schools, 19, 297-304.

Carvajal, H., & Weyand, K. (1986). Relationships between scores on Stanford-Binet IV and Wechsler Intelligence Scale for Children-Revised. Psychological Reports, 59, 963-966.

Chapman, J. W. and Boersma, F. J. (1979). Academic Self-Concept in Elementary Learning Disabled Children: A Study with the Student's Perception of Ability Scale. Psychology in the Schools, 16, 200-206.

Clariozo, H. and Bernard, R. (1981). Recategorized WISC-R Scores of Learning Disabled Children and Differential Diagnosis. Psychology in the Schools, 18, 5-12.

Cummins, J. R. & Das, J. P. (1980). Cognitive Processing, Academic Achievement and WISC-R Performance in EMR children. Journal of Consulting and Clinical Psychology, 48(6), 777-779.

Dudley-Marling, C. C., Kaufman, N. J. and Terver, S. G. (1981). WISC and WISC-R Profiles of Learning Disabled Children: A Review. Learning Disability Quarterly, 4, 307-319.

Ellis, H. C. and Hunt, R. R. (1983). Fundamentals of Human Memory and Cognition. (3rd ed.). Iowa: Wm C. Brown Company Publishers.

Epps, S., McGue, M., and Ysseldyke, J. (1982). Interjudge Agreement in Classifying Students as Learning Disabled. Psychology in the Schools, 19 (2), 209-220.

Epps, S., Ysseldyke, J. E., and Algozzine, B., (1985). An Analysis of the Conceptual Framework Underlying Definitions of Learning Disabilities. Journal of School Psychology, 23, 133-144.

Fagan,T. (1986a). Stanford-Binet Draws Board Attention. Communique, 15(1), 1.

Fagan, T. (1986b). Assembly withdraws Concern over Binet. Communique, 15 (3), 1.

Fisher, W. E. Wenck, L. S., Schurr, K. T., Ellen, A. S. (1985). The Moderating Influence of Gender, Intelligence, and Specific Achievement Deficiencies on the Bannatyne WISC-R Recategorization. Journal of Psychoeducational Assessment, 3, 245-255.

Fowler, P. C. (1986). Cognitive Differentiation of Learning-Disabled Children on the WISC-R: A Canonical Model of Achievement Correlates. Child Study Journal, 16, 25-37.

Galvin, G. A. (1981). Uses and Abuses of the WISC-R with the Learning Disabled. Journal of Learning Disabilities, 14, 326-329.

Gurney, P. G. (1986). Self-Esteem in the Classroom. School Psychology International, 7, 199-209.

Hale, R. L. (1978). The WISC-R as a Predictor of WRAT Performance. Psychology in the Schools, 15, 172-175.

Janzen, H. L. (1986). The Stanford-Binet IV: Subtest Analysis and Remediation Strategies. Unpublished Manuscript.

Janzen, H. L., Boersma, F. J., and Krausher, R. (1986). Report on the Binet IV. Unpublished Manuscript.

Janzen, H. L. Boersma, F. J., Fisk, R. and Chapman, J. (1983). Diagnosis and Predictability of Learning Disabilities. School Psychology International, 4, 129-140.

Jastak, S., & Wilkinson, G. S. (1984). The Wide Range Achievement Test Revised: Administration Manual. Wilmington, Deleware: Jastak Associates, Inc.

Karnes, F. A. & Brown, K. E. (1980). Factor Analysis of the WISC-R for the Gifted. Journal of Educational Psychology, 72, 197-199.

Kaufman, A. S. (1975). Factor Analysis of the WISC-R at 11 Age Levels between 6 1/2 and 161/2 Years. Journal of Consulting and Clinical Psychology, 43, 138-140.

Kaufman, A. S. (1976). A New Approach to the Interpretation of Test Scatter on the WISC-R. Journal of Learning Disabilities, 9, 160-168.

Kaufman, A. S. (1979). WISC-R Research: Implications for Interpretation. School Psychology Digest, 8, 5-27.

Kaufman, A. S. (1981). The WISC-R and Learning Disabilities Assessment: State of the Art. Journal of Learning Disabilities, 14(9), 520-527.

Kaufman, A. S. & McLean, J. E. (1986). K-ABC / WISC-R Factor Analysis for a Learning Disabled Population. Journal of Learning Disabilities, 19(3), 145-153.

Kaufman, A. & Van Hagan, J. (1977). Investigation of the WISC-R for use with retarded children: Correlation with the 1972 Stanford-Binet and Comparison of WISC and WISC-R Profiles. Psychology in the Schools, 14, 10-14.

Kavale, K. A. & Forness, S. R. (1984). A Meta-Analysis of the Validity of Wechsler Scale Profiles and Recategorizations: Patterns or Parodies? Learning Disability Quarterly, 7, 136-156.

Kavale, K. & Nye, Chad. (1981). Identification Criteria for Learning Disabilities: A Survey of the Research Literature. Learning Disability Quarterly, 4, 383-388.

Keith, T. Z. (1987). Assessment Research: An Assessment and Recommended Interventions. School Psychology Review, 16 (3), 276-289.

Kolligian, J., & Sternberg, R. (1987). Intelligence, Information Processing and Specific Learning Disabilities: A Triarchic Synthesis. Journal of Learning Disabilities, 20, 8-17.

Koppitz, E. M. (1981). The Bender Gestalt and VADS Test Performance of Learning Disabled Middle School Pupils. Journal of Learning Disabilities, 14, 96-98.

Leton, D. A., Meyamoto, L. K., and Ryckman, D. B. (1987). Psychometric Classifications of Learning Disabled Students. Psychology in the Schools, 24, 201-209.

Lubin, B., Larsen, R. M., and Matarazzo, J. D. (1984). Patterns of Psychological Test Usage in the United States: 1935-1982. American Psychologist, 39, 441-454.

Matuszek, P. (1985). Review of Wide Achievement Test, 1978 Edition In J.V. Mitchell (ED.), The Ninth Mental Measurements Yearbook. (Vol. II, 1334-1365).

McKinney, J. D. (1984). The Search for Subtypes of Specific Learning Disability. Journal of Learning Disabilities, 17(4), 43-49.

McManis, D. L., Figley, C., Richert, M and Fabre, T. (1978). Memory for Designs, Bender-Gestalt, Trail Making Test, and WISC-R Performance of Retarded and Adequate Readers. Perceptual and Motor Skills, 46, 443-450.

Meloff, R. (1987). The Stanford-Binet Intelligence Scale: Fourth Edition: A Descriptive Study. Unpublished Master's Thesis. University of Alberta, Edmonton, Alberta.

Miller, G. A. (1956). The Magical Number Seven, plus or minus two: Some Limits on Our Capacity for processing Information. Psychological Review, 63, 81-97.

Moore, D. W. & Wielan, O. P. (1981). WISC-R Scatter Indexes of Children Referred for Reading Diagnosis. Journal of Learning Disabilities, 14, 511-514.

Mueller, H. H., Dash, U. N., Matheson, D. W. and Short, R. H. (1984). WISC-R Subtest Patterning of Below Average, Average and Above Average IQ Children: A Meta-Analysis. The Alberta Journal of Educational Research, 30, 68--85.

Mueller, H. H., Dennis, S. S., Short, R. H., (1986). A Meta-Exploration of WISC-Factor Score Profiles as a Function of Diagnosis and Intellectual Level. Canadian Journal of School Psychology, 2, 21-43.

Naglieri, J. A. (1980). WISC-R Subtest Patterns for Learning Disabled and Mentally Retarded Children. Perceptual and Motor Skills, 51, 605-606.

Naglieri, J. A. (1982). Interpreting WISC-R and McCarthy Scatter: A Caution. Contemporary Educational Psychology, 7, 90-94.

Naglieri, J. A. (1985). Use of the WISC-R and K-ABC with Learning Disabled, Borderline Mentally Retarded, and Normal Children. Psychology in the Schools, 22, 133-141.

National Joint Committee on Learning Disabilities. (1987). Learning Disabilities: Issues on Definition. Journal of Learning Disabilities, 20(2), 107-108.

Osberg, D. W. (1986). New Tests: New Options. Communique. 15(3), 3.

Prasse, D. P., Siewert, J. C., Breen, M. J. (1983). An Analysis of Performance on Reading Subtests from the 1978 Wide Range Achievement Test and Woodcock Reading Mastery Test with the WISC-R for Learning Disabled and Regular Education Students. Journal of Learning Disabilities, 16, 458-461.

Rosenthal, R. & Rosnow, R. L. (1975). The Volunteer Subject. New York: John Wiley, 1975.

Rothlisberg, B. A. (1987). Comparing the Stanford-Binet, Fourth Edition to the WISC-R: A Concurrent Validity Study. Journal of School Psychology, 25, 193-196.

Ryckman, D. B. (1981). Searching for a WISC-R Profile for Learning Disabled Children: An Inappropriate Task? Journal of Learning Disabilities, 14, 508-527.

Sapp, G. L., Chissom, B. and Graham, E. (1985). Factor Analysis of the WISC-R for Gifted Students: A Replication and Comparison. Psychological Reports, 57, 947-951.

Sandovil, J. & Irvin, M. G. (1986). Stanford-Binet Intelligence Scale: Fourth Edition Technical Manual. CASP Today, 3-5.

Sattler, J. M. (1982). Assessment of Children's Intelligence and Special Abilities. (2nd ed.). Toronto: Allyn and Bacon, Inc.

Schakel, J. A. (1986). Cognitive Assessment of Preschool Children. School Psychology Review, 15 (2), 200-215.

Schiff, M. M. , Kaufman, A. S., and Kaufman, N. (1981). Scatter Analysis of WISC-R Profiles for Learning Disabled Children with Superior Intelligence. Journal of Learning Disabilities, 14, 400-405.

Scull, J. W. & Brand, L. H. (1980). The WRAT and the PIAT with Learning Disabled Children. Journal of Learning Disabilities, 13, 350-352.

Sexton, L. C. & Street, S. (1985). A Comparison of WISC-R Score Profiles of Students Placed in Special Education and Those Referred, But not Placed. Psychology in the Schools, 22, 142-145.

Sim, J. (1987). The Stanford-Binet IV: Recent Criticisms and Issues. The Alberta School Psychologist, 7 (1), 17-21.

Slate, J. R. (1986). A Reaction to the Revised Stanford-Binet Intelligence Scale: New does not Necessarily Mean Better. Communique, 15(1), 3.

Smith, C. R. (1983). Learning Disabilities. The Interaction of Learner, Task, and Setting. Toronto: Little, Brown and Company.

Smith, M. D. (1978). Stability of WISC-R Subtest Profiles for Learning Disabled Children. Psychology in the Schools, 15, 4-7.

Smith, M. D. , Coleman, J. M., Dolecki, P. R. and Davis, E. E. (1977). Recategorized WISC-R Scores of Learning Disabled Children. Journal of Learning Disabilities, 10(7), 48-54.

Smith, T. & Smith, B. (1986). The Relationship Between the WISC-R and WRAT-R for a Sample of Rural Referred Children. Psychology in the Schools, 23 (3), 252-254.

Snow, J. H, Cohen, M., and Holliman, W. B. (1985). Learning Disability Subgroups using Cluster Analysis of the WISC-R. Journal of Psychoeducational Assessment, 4, 391-397.

SPSS INC. (1983). SPSSX User's Guide. Chicago: SPSS Inc.

Stevenson, L. P. (1980). WISC-R Analysis: Implications for Diagnosis and Intervention. Journal of Learning Disabilities, 13, 346-349.

Strichart, S. S. & Love, E. (1979). WISC-R Performance of Children Referred to a University Center for Learning Disabilities. Psychology in the Schools, 16, 183-188.

Telzrow, C. F. (1987). New Stanford-Binet and Preschoolers. The Alberta School Psychologist, 7(1), 22.

Terman, L. M., & Merrill, M. A. (1937). Measuring Intelligence. Boston: Houghton Mifflin.

Terman, L. M. & Merrill, M. A. (1960). Stanford-Binet Intelligence Scale. Boston: Houghton Mifflin.

Terman, L. M., & Merrill, M. A. (1973). Stanford-Binet Intelligence Scale Manual for the Third Revision, Form L-M. Chicago: The Riverside Publishing Company.

Thorndike, R. L., Hagen, E. P. and Sattler, J. M. (1986a). Stanford-Binet Intelligence Scale: Fourth Edition, Guide for Administering and Scoring. Chicago: The Riverside Publishing Company.

Thorndike, R. L., Hagen, E. P. and Sattler, J. M. (1986b). Stanford-Binet Intelligence Scale: Fourth Edition, Technical Manual. Chicago: The Riverside Publishing Company.

Ullman, D. I., Kramer, J. J. & Schellenberg, R. P. (1985). WISC and WISC-R Range and Mean Deviation Indices of Scatter for a Clinic-Referred Population. Psychological Reports, 66, 841-842.

Vance, H. B. & Singer, M. G. (1979). Recategorization of the WISC-R Subtest Scaled scores for Learning Disabled Children. Journal of Learning Disabilities, 12(8), 487-491.

Vernon, P. E. (1987). The Demise of the Stanford-Binet Scale. Canadian Psychology, 28(3), 251-258.

Wechsler, D. (1974). Manual for the Wechsler Intelligence Scale for Children-Revised. New York: Psychological Corporation.

Westling, D. L. (1986). Introduction to Mental Retardation. Englewood Cliffs, New Jersey: Prentice Hall, Inc.

Wilde, W. D. & Sillito, M. T. (1986). Evaluation Report of the Challenge Program, County of Strathcona #20. Alberta: Alberta Education.

Wright, D., & DeMers, S.T. (1982). Comparison of the Relationship Between Two Measures of Visual-Motor Coordination and Academic Achievement. Psychology in the Schools, 19, 473-477.

Yewchuk, C. R. (1986). Identification of Gifted / Learning Disabled Children. School Psychology International, 7, 61-68.

Ysseldyke, J. E. , Algozzine, B. , Shinn, M. R. , McGue, M. (1982). Similarities and Differences Between Low Achievers and Students Classified Learning Disabled. The Journal of Special Education, 16, 73-85.

APPENDIX A
DISTRIBUTION OF THE SAMPLE BY DEMOGRAPHIC CHARACTERISTICS

TABLES 1 - VI

TABLE I
DISTRIBUTION OF THE SAMPLE BY AGE

AGE	N	%	AGE	N	%
2-0 to 2-11	1	.30	10-0 to 10-11	28	8.6
3-0 to 3-11	10	3.1	11-0 to 11-11	19	5.8
4-0 to 4-11	20	6.1	12-0 to 12-11	21	6.4
5-0 to 5-11	47	14.4	13-0 to 13-11	17	5.2
6-0 to 6-11	26	8.0	14-0 to 14-11	15	4.6
7-0 to 7-11	29	8.9	15-0 to 15-11	11	3.4
8-0 to 8-11	16	4.9	16-0 to 16-11	21	6.5
9-0 to 9-11	28	8.6	17-0 to 17-11	6	1.8
			18-0 to 27-03	11	3.4
TOTAL N =326			MEAN AGE = 9.6 Years		

TABLE II
DISTRIBUTION OF THE SAMPLE BY GENDER

SEX	NUMBER	PERCENT
Male	186	57.1
Female	139	42.6
TOTAL	326	100.0

TABLE III
DISTRIBUTION OF THE SAMPLE BY ETHNICITY

ETHNICITY	NUMBER	PERCENT
Caucasian	273	83.7
Native	23	7.1
Asian	14	4.3
Hispanic	11	3.4
Other	5	1.5
TOTAL	326	100.0

TABLE IV
DISTRIBUTION OF THE SAMPLE BY LOCALE

LOCALE	NUMBER	PERCENT
Large Cities		
299,999 or larger	212	65.0
100,000 to 299,999	1	0.3
Cities		
25,000 to 99,999	26	8.0
Small Towns		
2500 to 24,999	52	16.0
Rural Areas		
Less than 2500	35	10.7
TOTAL	326	100.0

TABLE V
DISTRIBUTION OF THE SAMPLE BY PARENTAL OCCUPATION

OCCUPATION	NUMBER	PERCENT
Managerial/Professional	89	27.3
Technical/Sales	80	24.5
Service Occupations	81	24.8
Farming/Forestry	20	6.1
Precision Production	9	2.8
Operators/Fabricators	11	3.4
Unemployed	9	2.8
Unknown	19	5.8
Other	8	2.5
TOTAL	326	100.0

TABLE VI
DISTRIBUTION OF THE SAMPLE BY PARENTAL EDUCATION

EDUCATION	NUMBER	PERCENT
College Graduate or Beyond	84	25.8
1 - 3 years College	65	19.9
HS Grad	65	19.9
< HS	23	7.1
Unknown	89	27.3
TOTAL	326	100.0

TABLE VII
DISTRIBUTION OF THE SAMPLE BY PARENTAL MARITAL STATUS

MARITAL STATUS	NUMBER	PERCENT
Married	277	85.0
Divorced	7	2.2
Separated	21	6.4
Single	3	.9
One Parent Deceased	4	1.2
Parent Remarried	4	1.2
Unknown	10	3.1
TOTAL	326	100.0

APPENDIX B

PSYCHOLOGICAL TESTING CENTER

FACULTY OF EDUCATION, CLINICAL SERVICES

REPORT FORWARDING FORM

Table VIII

TABLE VIII
REPORT FORWARDING FORM*

REASON FOR REFERRAL: (Check one or several)

_____Readiness for Kindergarten (4-5)
_____Pre-school Assessment (2-4)
_____Learning Disabilities (Child)
_____Dyslexia (Child or Adult)
_____School Learning Problems
_____Behavior Problems
_____Developmental Delay/Maturation
_____Social/Emotional Problems(General)
_____Self-Referral - General Interest
_____Career Choices/Planning
_____Personality/AbilityInterest
_____Reading/Language Problems
_____Neuropsychological
_____Child Custody
_____Guaranteed Income (Adult)
_____Giftedness/Enrichment
_____Parental Interest (In Child's Ability)
_____Special Class or Program Placement
_____Special School Placement
_____Mental Retardation
_____Sensory Impairment (Hearing, Vision,speech)

**Reason for Referral,* used for this study is only part of the complete "Report Forwarding Form" used in the Faculty of Education, Clinical Services.

APPENDIX C
ONEWAY ANALYSIS OF VARIANCE
SBIV, WRAT-R, PASS AND VMI SCORES
TABLES IX - XXXV

TABLE IX
ONEWAY ANALYSIS OF VARIANCE
SBIV COMPOSITE IQ

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between groups	3	96203.9072	32067.9691	282.8533	.0000
Within groups	262	29703.5514	113.3723		
Total	265	125907.4586			

TABLE X
ONEWAY ANALYSIS OF VARIANCE
SBIV VERBAL REASONING IQ

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between groups	3	70976.7126	23658.9042	168.5625	
Within groups	262	36773.4980	140.3569		
Total	265	107750.2105			

TABLE XI
ONEWAY ANALYSIS OF VARIANCE
SBIV ABSTRACT/VISUAL REASONING IQ

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	74208.3099	24736.1033	168.2169	.0000
Within Groups	261	38379.7505	147.0489		
Total	264	112588.0604			

TABLE XII
ONEWAY ANALYSIS OF VARIANCE
SBIV QUANTITATIVE REASONING IQ

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	67984.1973	22661.3991	170.4063	.0000
Within Groups	259	34442.9814	132.9845		
Total	262	102427.1787			

TABLE XIII
ONEWAY ANALYSIS OF VARIANCE
SBIV SHORT-TERM MEMORY IQ

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	73858.4941	24619.4980	128.2945	.0000
Within Groups	262	50277.3443	191.8983		
Total	265	124135.8383			

TABLE XIV
ONEWAY ANALYSIS OF VARIANCE
SBIV VOCABULARY

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	14804.8290	4934.9430	138.5771	.0000
Within Groups	262	9330.2236	35.6115		
Total	265	24125.0526			

TABLE XV
ONEWAY ANALYSIS OF VARIANCE
SBIV COMPREHENSION

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	13867.8403	4622.6134	93.5227	.0000
Within Groups	257	12702.9183	49.4277		
Total	260	26570.7586			

TABLE XVI
ONEWAY ANALYSIS OF VARIANCE
SBIV ABSURDITIES

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	10235.3152	3411.7717	76.4911	.0000
Within Groups	230	10258.8087	44.6035		
Total	233	20494.1239			

TABLE XVII
ONEWAY ANALYSIS OF VARIANCE
SBIV PATTERN ANALYSIS

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	12382.9255	4127.6452	82.5446	.0000
Within Groups	261	13051.3060	50.0050		
Total	264	25434.2415			

TABLE XVIII
ONEWAY ANALYSIS OF VARIANCE
SBIV COPYING

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	14031.9571	4677.3190	114.7941	.0000
Within Groups	259	10553.0315	40.7453		
Total	262	24584.9886			

TABLE XIX
ONEWAY ANALYSIS OF VARIANCE
SBIV MATRICES

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	14031.9571	4677.3190	114.7941	.0000
Within Groups	259	10553.0315	40.7453		
Total	262	24584.9886			

TABLE XX
ONEWAY ANALYSIS OF VARIANCE
SBIV NUMBER SERIES

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	7172.2345	2390.7448	92.3356	.0000
Within Groups	147	3806.1099	25.8919		
Total	150	10978.3444			

TABLE XXI
ONEWAY ANALYSIS OF VARIANCE
SBIV BEAD MEMORY

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	12581.3694	4193.7898	58.9526	.0000
Within Groups	260	18495.9602	71.1383		
Total	263	31077.3295			

TABLE XXII
ONEWAY ANALYSIS OF VARIANCE
SBIV MEMORY FOR SENTENCES

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	15546.1524	5182.0508	94.7521	.0000
Within Groups	258	14110.1720	54.6906		
Total	261	29656.3244			

TABLE XXIII
ONEWAY ANALYSIS OF VARIANCE
SBIV MEMORY FOR DIGITS

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	7401.0558	2467.0186	57.7635	.0000
Within Groups	209	8926.1649	42.7089		
Total	212	16327.2207			

TABLE XXIV
ONEWAY ANALYSIS OF VARIANCE
SBIV MEMORY FOR OBJECTS

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	4725.7395	1575.2465	34.9049	.0000
Within Groups	190	8574.6368	45.1297		
Total	193	13300.3763			

TABLE XXV
ONEWAY ANALYSIS OF VARIANCE
WRAT-R READING

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	41895.1150	13965.0383	69.0234	.0000
Within Groups	206	41678.5803	202.3232		
Total	209	83573.6952			

TABLE XXVI
ONEWAY ANALYSIS OF VARIANCE
WRAT-R SPELLING

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	41074.6740	13961.5580	73.7775	.0000
Within Groups	207	38414.8710	185.5791		
Total	210	79489.5450			

TABLE XXVII
ONEWAY ANALYSIS OF VARIANCE
WRAT-R ARITHMETIC

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	53840.1938	17946.7313	102.7795	.0000
Within Groups	207	36145.0764	174.6139		
Total	210	89985.2701			

TABLE XXVIII
ONEWAY ANALYSIS OF VARIANCE
PASS FULL SCALE

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	2423.6359	807.8786	5.9853	.0000
Within Groups	116	15657.2891	134.9766		
Total	119	18080.9250			

TABLE XXIX
ONEWAY ANALYSIS OF VARIANCE
PASS GENERAL ABILITY

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	150.4514	50.1505	5.8971	.0009
Within Groups	115	977.9856	8.5042		
Total	118	1128.4370			

TABLE XXX
ONEWAY ANALYSIS OF VARIANCE
PASS ARITHMETIC

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	77.6816	25.8939	3.3394	.0218
Within Groups	115	891.7134	7.7540		
Total	118	969.3950			

TABLE XXXI
ONEWAY ANALYSIS OF VARIANCE
PASS SCHOOL SATISFACTION

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	86.3239	28.7746	3.7709	.0126
Within Groups	115	877.5417	7.6308		
Total	118	963.8655			

TABLE XXXII
ONEWAY ANALYSIS OF VARIANCE
PASS READING/SPELLING

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	154.9271	51.6424	4.7666	.0036
Within Groups	115	1245.9468	10.8343		
Total	118	1400.8739			

TABLE XXXIII
ONEWAY ANALYSIS OF VARIANCE
PASS PENMANSHIP/NEATNESS

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	17.0314	5.6771	.6644	.5756
Within Groups	115	982.6660	8.5449		
Total	118	999.6975			

TABLE XXXIV
ONEWAY ANALYSIS OF VARIANCE
PASS CONFIDENCE

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	69.1164	23.0388	4.7328	.0038
Within Groups	115	559.8079	4.8679		
Total	118	628.9244			

TABLE XXXV
ONEWAY ANALYSIS OF VARIANCE
VMI

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB
Between Groups	3	595.3227	198.4409	24.8934	.0000
Within Groups	88	701.5034	7.9716		
Total	91	1296.8261			

B45573